EDGEWOOD
CHEMICAL BIOLOGICAL CENTER
U.S. ARMY SOLDIER AND BIOLOGICAL CHEMICAL COMMAND

ECBC-SP-001

PROCEEDINGS OF THE FIRST JOINT SERVICES WORKSHOP ON BIOLOGICAL MASS SPECTROMETRY

28-30 July 1997
Baltimore, Maryland

S. Randolph Long
Thaiya Krishnamurthy
A. Peter Snyder

RESEARCH AND TECHNOLOGY DIRECTORATE

January 1999

Aberdeen Proving Ground, MD 21010-5424

Disclaimer

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorizing documents.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave Blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	1999 January	Final; 1997 July

4. TITLE AND SUBTITLE

Proceedings of the First Joint Services Workshop on Biological Mass Spectrometry, 28-30 July 1997, Baltimore, Maryland

5. FUNDING NUMBERS

PR-643884/ACA4

6. AUTHOR(S)

Long, S. Randolph; Krishnamurthy, Thaiya; and Snyder, A. Peter

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

DIR, ECBC,* ATTN: AMSSB-RRT-E, APG, MD 21010-5424

8. PERFORMING ORGANIZATION REPORT NUMBER

ECBC-SP-001

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

*When this workshop was conducted, ECBC was known as the U.S. Army Edgewood Research, Development and Engineering Center (ERDEC).

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The first Joint Services Workshop on Biological Mass Spectrometry was held 28-30 July 1997, at the Hyatt Regency, Baltimore, MD. The workshop was organized by the U.S. Army Edgewood Research, Development, and Engineering Center (ERDEC)* and co-sponsored by the U.S. Navy and U.S. Air Force. The overall objective of the workshop was to bring together experts and practitioners for a review of the state-of-the-art in biological mass spectrometry and related technologies to serve as the basis for the generation of a program strategy for the development of next-generation biological mass spectrometers. Such mass spectrometer systems will have significant impact not only on chemical/biological defense but also on environmental and food monitoring, the health industry, and various research and industrial endeavors. The workshop addressed a range of issues related to the development of automated, integrated mass spectrometers based on leading edge methodologies for the characterization, analysis, and identification of biological materials. Particular attention was given to applications of electrospray ionization (ESI) and matrix-assisted laser desorption ionization (MALDI), interface of separation methods, biological sample handling/processing, and potential biomarker classes. These Proceedings compile summaries of a majority of technical presentations made during the conference portion of the workshop.

14. SUBJECT TERMS

Biological detection
Biological mass spectrometry

15. NUMBER OF PAGES

209

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	**UNCLASSIFIED**	**UNCLASSIFIED**	**UL**

Blank

PREFACE

The work described in this report was authorized under Project No. 643884/ACA4. This work was started and completed in July 1997.

The use of either trade or manufacturers' names in this report does not constitute an official endorsement of any commercial products. This report may not be cited for purposes of advertisement.

This report has been approved for public release. Registered users should request additional copies from the Defense Technical Information Center; unregistered users should direct such requests to the National Technical Information Service.

Blank

TABLE OF CONTENTS

SESSION I

MODERATOR: Dr. Thaiya Krishnamurthy
U.S. Army Edgewood Research, Development and Engineering Center*
Aberdeen Proving Ground, MD

SESSION II

MODERATOR: Dr. A. Peter Snyder
U.S. Army Edgewood Research, Development and Engineering Center
Aberdeen Proving Ground, MD

* Now known as the U.S. Army Edgewood Chemical Biological Center.

TABLE OF CONTENTS

TABLE OF CONTENTS

SESSION IV

MODERATOR: Dr. Ngai M. Wong
U.S. Air Force Armstrong Laboratory
Aberdeen Proving Ground, MD

TABLE OF CONTENTS

SESSION V

MODERATOR: Dr. Harold Guard
U.S. Office of Naval Research
Arlington, VA

TABLE OF CONTENTS

SESSION VI

MODERATOR: Dr. S. Randolph Long
U.S. Army Edgewood Research, Development and Engineering Center
Aberdeen Proving Ground, MD

POSTER PRESENTATIONS

MODERATOR: Dr. Jhobe Steadman
Geo-Centers, Inc.
Aberdeen Proving Ground, MD

TABLE OF CONTENTS

BIOLOGICAL MASS SPECTROMETRY [1]

Catherine E. Costello
Boston University School of Medicine
Mass Spectrometry Resource, Dept. of Biophysics
715 Albany Street, R-806
Boston, MA 02118-2394 USA

Electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI) combine high sensitivity with wide mass range and now permit the analysis of subpicomole (indeed, attomole) amounts of materials having molecular weights up to megadaltons. These central techniques are user-friendly, an enormously powerful trait that has encouraged the entry of many from adjacent fields into the prior preserve of chemists and physicists. Such access, combined with developments in sample delivery techniques, analyzer options, software for instrument control and data interpretation, and internet communications, has resulted in utilization of mass spectrometry among a very broad range of scientists. This workshop surveyed the *status quo* in the summer of 1997. We began with a summary description of terminology, methods and approaches. Examples drawn largely from studies conducted in our laboratory were used as illustrations. Our research focuses on development and application of mass spectral methods for glycobiology, a field that deals with crucial but quite incompletely-understood areas, *e.g.*, immunology, growth and development, transplantation, carcinogenesis, parasitology, intra- and intercellular signalling and protein transport; it probes structure-activity relationships pertinent to biology and medicine and requires all the sensitivity, specificity and mass range that these emerging methods can deliver.

Our approaches and the current state of the field for biological mass spectrometry are summarized in a review now in press for *Biophysical Chemistry*. The galley proofs of this review "Time, Life ..... and Mass Spectrometry" are included here as fair representation of the talk presented at the Workshop.

[1]Research support has been provided by NIH National Center for Research Resources Grant Nos. P41 RR10888 and S10 RR10493, Thermo BioAnalysis Corp. and AutoImmune, Inc.

ELSEVIER

Biophysical Chemistry 2301 (1997) xxx C

Biophysical Chemistry

Time, life . . . and mass spectrometry New techniques to address biological questions

Catherine E. Costello *

Mass Spectrometry Resource, Dept. of Biophysics, Boston Univ. School of Medicine, Boston USA

Received 9 January 1997; revised 21 February 1997; accepted 21 February 1997

Abstract

Within the last ten years, startling new developments in two ionization methods—matrix-assisted laser desorption (MALDI) and electrospray (ESI)—have been described by Karas et al. [M. Karas, D. Bachmann, U. Bahr, F. Hillenkamp, Int. J. Mass Spectrom. Ion Proc., 78 (1987) 53.] and by Fenn et al. [J.B. Fenn, M. Mann, C.K. Meng, S.F. Wong, C.M. Whitehouse, Science, 246 (1989) 64.], respectively. Their work demonstrated that these techniques, under appropriate experimental conditions, have high sensitivity and wide mass range, extending to hundreds of thousands of daltons and beyond, and thus can be extremely effective for the study of biopolymers. The result has been a revolution in the way that mass spectrometry experiments are carried out, a widening of the range of investigators who employ mass spectrometry in their own laboratories and a penetration of mass spectrometry into the investigation of biological phenomena that exceeds any previous expectations. Progress in improving mass spectral ionization and mass analysis methods and in interpreting and understanding the spectra is actively being pursued and exploited in many laboratories, to capitalize even further upon these advances. The results should facilitate understanding of structure–activity relationships pertinent to biology and medicine. In our laboratory, the focus of research is on oligosaccharide and glycoconjugate structural determinations, and on the improvement of methods for these important classes of compounds that relate to development, immune response, signalling, lipid and protein transport and disease. Representative examples of applications of MALDI and ESI mass spectrometry to these and other biological questions are provided herein. © 1997 Elsevier Science B.V.

Keywords: Mass spectrometry; Electrospray ionization; Matrix-assisted laser desorption/ionization; Glycoproteins; Glycolipids; Oligosaccharides; Muramyl peptides; Proteins

1. Introduction

During the last decade, dramatic progress has been made in methods for the structural determination and analysis of biological molecules, largely through exploitation of the advantages of two ionization methods that increase both sensitivity and accessible mass range: matrix-assisted laser desorption/ionization (MALDI) [1] and electrospray ionization (ESI) [2]. In conjunction with the availability of these ionization methods, advances in instrumentation, data processing and information retrieval have also been rapid and significant. It is now possible to

* Corresponding author. Mass Spectrometry Resource, Dept. of Biophysics, 80 E. Concord Street, R-806, Boston, MA 02118-2394. Tel.: +1-617-638-6490; fax: +1-617-638-6491; e-mail: cecmsms@bu.edu.

PII S0301-4622(97)00033-1

determine molecular weight profiles with low femtomole sensitivity and to use tandem mass spectrometry (MS/MS) to elucidate structural details at the low- or sub-picomole level for individual compounds or the components of mixtures. The spectra can be compared to internationally available databases, or interpreted with the guidance of rules being developed for standard sets of compounds. Biophysical data can be obtained regarding interactions with ligands and cations, ideally under conditions that approach or mimic the physiological norm. As the reach of mass spectrometry is extended to higher molecular weight compounds—now up to hundreds of thousands of daltons and beyond—and to trace components in biological systems, new methods for sample isolation and handling are required. Concomitantly, a new generation of 'user-friendly' instruments is being developed to bring the techniques directly into the workplaces of biological and clinical investigators.

The range of questions whose answers are now amenable to mass spectral pursuit is illustrated here with a few examples from our laboratory. Some collaborative studies involve proteins, peptide receptors and oligonucleotides, but our primary research emphasis is placed on oligosaccharides and glycoconjugates. We seek to improve the mass spectral sensitivity for these compounds and to increase the information content in the spectra, in order to facilitate investigations of their structure–activity relationships in biological systems. Because of the complexity of the structures and the regular occurrence of mixtures of related compounds with varying activities, the analytical problems for carbohydrates exceed those of linear biopolymers such as proteins and oligonucleotides. The challenge must be met, however, because increasing evidence demonstrates that oligosaccharides and their conjugates play many diverse and important roles [3]. Our areas of interest include immunology, nervous system growth and development, transplantation, carcinogenesis, infectious diseases, and parasitology.

Matrix-assisted laser desorption / ionization time-of-flight mass spectrometry (MALDI-TOF MS) is illustrated in Fig. 1(a). The sample is mixed with a thousand-fold excess of a material which absorbs at the wavelength of the laser (337 nm for the nitrogen laser employed for most of the examples described herein). By mechanisms that are as yet incompletely understood, the energy transfer to the matrix results in desorption of both matrix and analyte from the surface and the formation of ions within the plume. It is convenient to measure the spectra of the desorbed ions with a time-of-flight mass analyzer, because this device is immediately compatible with the pulsed generation of high-energy ions and the mass range is theoretically unlimited. Linear mode spectra are acquired when the ion currents are measured at Detector 1. The use of a reflectron, as illustrated here with signal recorded at Detector 2, increases the resolution through energy focusing of the ions in a mirror [4]. When set to transmit the ions that have full energy (stable molecular ions and fragments

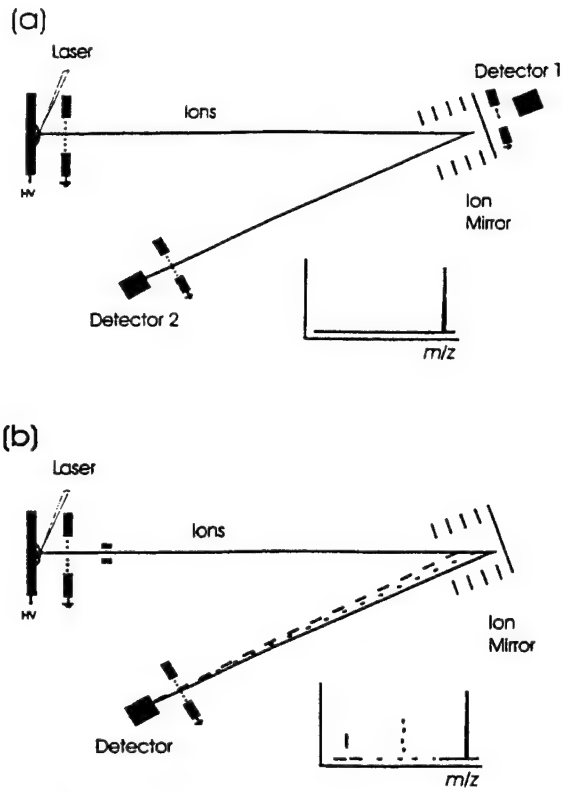


Fig. 1. (a) Scheme of matrix-assisted laser ionization/desorption and mass analysis in a reflectron time-of-flight mass spectrometer. Not to scale: the ion flight path is usually 1–3 m, whereas dimensions in the ion source and detector regions are less than a few cm. (b) Ion trajectories in a reflectron time-of-flight mass spectrometer as a function of reflector voltages. Ion transmission pathways during various stages of the post-source decay experiment are indicated.

generated from very unstable molecular ions that underwent 'prompt' fragmentation before acceleration), the reflectron rejects lower energy fragmentations that have arisen during flight. In an experiment that utilizes stepwise decrease of the reflector potential, these lower energy ions may be transmitted to the detector to capture the structural information contained in the fragmentation pattern, as illustrated in Fig. 1(b). This type of data acquisition is now designated as the post-source decay (PSD) mode [5–7] and corresponds to the multistage (MS/MS) measurement of the metastable ion spectrum on other instruments. Sensitivity for PSD is in the sub-picomole range. The performance of even a simple linear device may be significantly improved by allowing a brief (nanoseconds) pause between ion generation and acceleration (time-delayed extraction) [8–10]. The delay results in extraction of ions within a narrower energy range. Other types of analyzers may also be utilized, with varying degrees of cost and requirements for operator expertise. Double-focusing instruments with integrating detectors and ion traps have been used (*mutatis mutandi*) in conjunction with MALDI, and have yielded promising results.

Electrospray ionization mass spectrometry (ESI-MS) and tandem mass spectrometry (ESI-MS/MS) are illustrated in Fig. 2(a) and (b), respectively. In this mode, the sample is introduced in a flowing stream, driven from a syringe pump or eluting from an HPLC or capillary zone electrophoresis system. Maximum sensitivity is observed when the sample solution is admitted directly to the ion source through an extremely finely drawn tip (1–3 μM) that delivers 1 μL over about 30 min [11]. The flow contains an electrolyte, such as sodium acetate, and may include other additives, buffers or 'electron traps', e.g., chlorinated hydrocarbons. As the flow discharges from the tiny nozzle held at high voltage with respect to the ion source lens, droplets emerge bearing the electrolyte on the surface and carrying the sample within. As solvent is stripped away in the vacuum, the droplet size decreases, and sample-rich particles result, often bearing high numbers of charges. The charged molecular ions enter the analyzer region with little excess energy and are therefore quite stable; ESI mass spectra contain mostly molecular ions in a distribution of charge states. Fragmentation may be induced by manipulation of

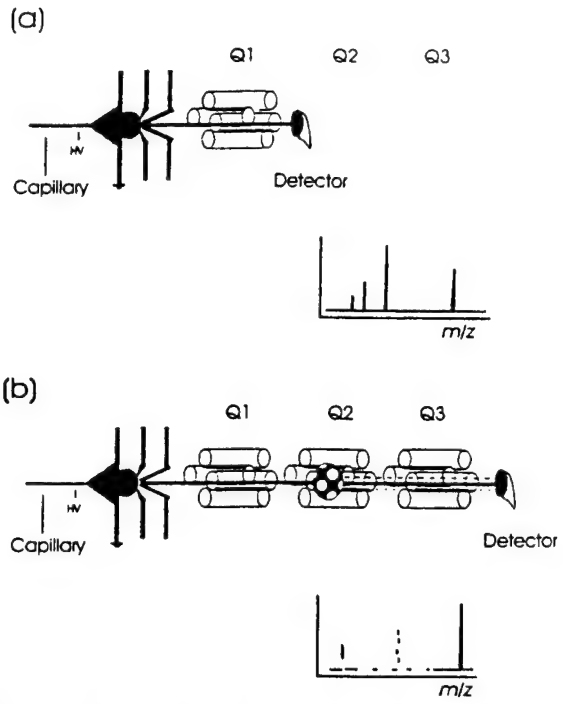


Fig. 2. (a) Electrospray ionization with mass separation in a single-stage quadrupole MS. (b) Ionization and mass separation during an ESI-MS/MS experiment in a triple quadrupole mass spectrometer. The central quadrupole, Q2, serves as the collision cell. Precursor ions selected by Q1 are decomposed in Q2; the product ions are separated in Q3 and detected.

the ion source repeller lens voltage, or by collision of mass-selected ions in the region between MS-1 (Q1) and MS-2 (Q3). The latter approach to ESI-MS/MS is usually preferable, because it allows precursor ion selection and better control over the decomposition process. Although single and triple quadrupole instruments dominate the ESI literature, ESI may also be performed with ion traps, time-of-flight and multisector mass analyzers. With ion traps, whether quadrupolar or Fourier transform ion cyclotron resonance (FT-ICR), multiple stages of mass selection and decomposition are possible, as well as studies of ion-molecule reactions. ESI investigations of noncovalent complexes (e.g., receptor–ligand binding) and differential rates of deuterium exchange [12,13] are particularly relevant for biological studies.

2. Experimental

2.1. MALDI-TOF MS

MALDI-TOF MS was carried out using a Finnigan MAT (now Thermo BioAnalysis, Franklin, MA) Vision 2000 reflectron time-of-flight instrument operated at ±30 kV (linear mode) or ±5–10 kV (reflectron and PSD modes). In the ultraviolet (UV) range, a nitrogen laser (Laser Science, Newton, MA) was used to generate the primary beam, at 337 nm, with 3–5 ns pulse width. In the infrared (IR) range, an Er-YAG laser (finYAG 20-2, Spektrum, Berlin) generated a beam at 2.94 μm, with 50–100 ns pulse width. PSD mass spectra were acquired in 10–15 segments, by stepping down the reflector voltage and were reassembled using the instrument's software package. Samples were applied to the stainless steel target as 10^{-5}–10^{-6} M solutions (1 μL) mixed with an equal volume of the 10^{-4} M matrix solution (usually 2,5-dihydroxybenzoic acid for UV-MALDI, succinic acid for IR-MALDI). The applied drop was allowed to air dry and crystallize before the probe was introduced into the mass spectrometer ion source.

2.2. ESI-MS and ESI-MS / MS

ESI-MS and ESI-MS/MS experiments were carried out with a VG (now Micromass, Beverly, MA) Quattro II triple quadrupole mass spectrometer or a Finnigan (now Thermo BioQuest, Sunnyvale, CA) TSQ 700 triple quadrupole tandem mass spectrometer. In either case, collision-induced decomposition took place with Ar (2 mTorr) in the second (rf-only) quadrupole. Sample solutions (ca. 0.1 pmol/μL in methanol/water (6:4 v/v) containing 0.25 mM NaOH) were introduced with a Harvard syringe pump, at a flow rate of 0.85 μL/min.

2.3. Sample preparation

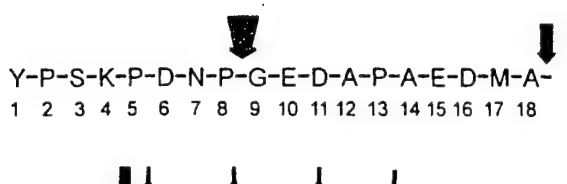
For MALDI-TOF MS, the 2,5-dihydroxybenzoic acid (Aldrich, Milwaukee, WI) matrix was recrystallized or subjected to ion exchange to control cation content. Native samples of oligosaccharides, glycoconjugates, proteins, peptides or oligonucleotides were desalted by passage through a C-18 column (HPLC or Sep-Pak®), microdialysis on a floating membrane [14,15] or ion-exchange [16] until salt content was eliminated or minimized, as judged from the mass spectra. For both MALDI-TOF MS and ESI-MS, oligosaccharides and glycolipids were permethylated using the solid NaOH/methyl iodide procedure of Ciucanu and Kerek [17], adapted for microscale [18], or peracetylated in capillaries using a gas/solid reaction [19].

3. Results and discussion

We move easily back and forth between MALDI and ESI techniques, and frequently use both approaches during a single analysis. This provides information on the unique properties (and potential biases) of the two methods, and allows the selection of optimum conditions for determinations related to samples which are available only in very small amounts because of their biological occurrence at low levels in microscopic tissues or extraordinary requirements for isolation and purification that preclude preparation of larger quantities. These may still be contaminated and may require microscale cleanup, derivatization or degradation to complete the determinations. As an illustration of the approaches, a few recent examples from our own laboratory and collaborative projects are presented here.

3.1. Studies of neuropeptides

Biologically active neuropeptides are generated from much larger preprohormones in response to various stimuli; the initial products from the preprohormones may undergo several transformations to the active form, before they are degraded and inactivated [20]. The timing and course of these processes are controlled by normal and pathological conditions. The pattern of neuropeptides found in a given tissue at any moment thus represents a window on a number of interrelated metabolic pathways. While radioimmunoassay (RIA) often is seen as a convenient method for estimating neuropeptide levels, it can suffer from lack of specificity when presented with a series of closely related compounds, and loss of utility when the recognized functionality is substantially modified or deleted [21]. With the advent of MALDI and ESI techniques, MS sensitivity equals or exceeds that of radioimmunoassay [22–24], provides specific molecular weight information and can



Scheme 1. Structure of neuropeptide Y with arrows indicating the cleavage sites observed for incubation in human CSF. Reprinted (with permission) from Ref. [25].

furnish structural details when required.

3.2. Neuropeptides in human cerebrospinal fluid (CSF)

The fact that neuropeptides may be analyzed by MALDI-TOF MS with little sample preparation has been demonstrated [22,23]. One of the most abundant mammalian CNS neuropeptides is neuropeptide Y (NPY), a 36-residue peptide with a C-terminal amide (Scheme 1) [26,27]. Neuropeptide Y has been shown to control, through hypothalamic mechanisms, the release of several pituitary hormones, to stimulate the appetite, especially for carbohydrates [28], and to affect emotions [29–31]. There is evidence for the existence of several subtypes of neuropeptide Y receptors which individually have a specific requirement for either the intact peptide or a truncated fragment [32]. Information on complete peptide processing patterns has not yet been available. To address this need, in collaboration with R. Ekman et al. [25], Univ. Göteborg, we recently followed and compared the processing of neuropeptide Y in the CSF of normals and depressed patients. CSF samples lyophilized in phosphate buffer (10 mM pH 7.0) were reconstituted to their original volumes with water, spiked with NPY solution to a final concentration of 2 fmol → 1 pmol/μL and incubated at 37°C. For MALDI-TOF MS kinetic profiles, 1-μL aliquots were withdrawn and mixed with 5–10 volumes of matrix solution; spectra were acquired for 1 μL of the resulting mixture. We found that the major products were formed by cleavage of single peptide bonds in NPY to yield N-terminal or C-terminal fragments (Scheme 1). Although the product distribution was fairly constant among individuals, the degradation rates varied sharply. Within the small sample examined thus far, depressed patients showed rates faster than normals. Further MALDI-TOF MS studies will explore the apparent correlation, define processing pathways in different groups or in individuals over time, and may lead to clinical assays and a better understanding of the biochemical basis of some types of depression.

3.3. Glycoforms on rat pineal gland glycopeptide

The sensitivity of MALDI-TOF MS and ESI-MS are such that studies may be carried out on single cells or small biopsy samples. In many cases, the lowest practical limit is one of sample-handling rather than mass spectral sensitivity. Our collaborators in the Vrie Universität, Amsterdam, have demonstrated that peptides in single snail neurons can be profiled by mixing individual cells with the MALDI matrix, applying the mixture to the MS target, and irradiating with a nitrogen laser [23]. In a recent study, they undertook profiling the peptides in the neurointermediate lobe of the rat pituitary gland under normal conditions and compared the pattern to that observed after salt-loading [33]. While some peptide levels remained fairly constant, the levels of the suite of peptides that are released from the preprohormones propressophysin and prooxyphysin were distinctly lowered. Included in this group was a component whose partial amino acid sequence (AREQS?ATQL...) was determined by microscale Edman degradation and corresponded to the C-terminal peptide [A(129)REQSNATQL...Y(168)] of propressophysin (CPP), but whose apparent molecular weight, 5529 daltons, indicated a posttranslational modification. The peptide sequence contained a potential *N*-glycosylation site and the Edman results suggested that the Asn residue was modified. We received an HPLC fraction containing this candidate glycopeptide obtained from extraction of three cells, an amount estimated to be about 200 pmol. The sample was examined by MALDI-TOF MS and a portion (about 150 pmol) was treated with *N*-glycanase to release the carbohydrate. The carbohydrate was purified by passage through a C18 Sep-pak® and permethylated by a modification of

the solid NaOH/methyl iodide method described by Ciucanu and Kerek [17] and Linsley et al. [18]. The permethylated oligosaccharide fraction was dissolved in methanol/water (6:4 v/v) containing 0.25 mM NaOH and analyzed by ESI-MS and ESI-MS/MS. This procedure has been developed in our laboratory for profiling glycoprotein oligosaccharides at high sensitivity [34].

In the ESI mass spectrum, Fig. 3(a), the doubly-charged sodiated molecular ion $[M + 2Na]^{2+}$ of the major component was observed at m/z 1052.6, and a minor component had $[M + 2Na]^{2+}$ at m/z 1241.0. These correspond to permethylated molecular weights of 2059.3 and 2437.7, respectively, and imply native molecular weights of 1666.5 and 1974.5 for the released oligosaccharides. The molecular weights calculated for the terminal CPP peptide with these glycoforms would be 5930 and 6238 daltons. These results were consistent with the MALDI-TOF mass spectrum, wherein the observed $[M + H]^+$ values were m/z 5930 and 6239. The mass spectral measurements thus suggested the oligosaccharide compositions $Hex_3HexNAc_5Fuc$ and $Hex_4HexNAc_5Fuc_2$.

In order to investigate details of the peptide glycosylation, the $[M + 2Na]^{2+}$ ion from each released and permethylated oligosaccharide was analyzed by ESI-MS/MS with collision-induced disso-

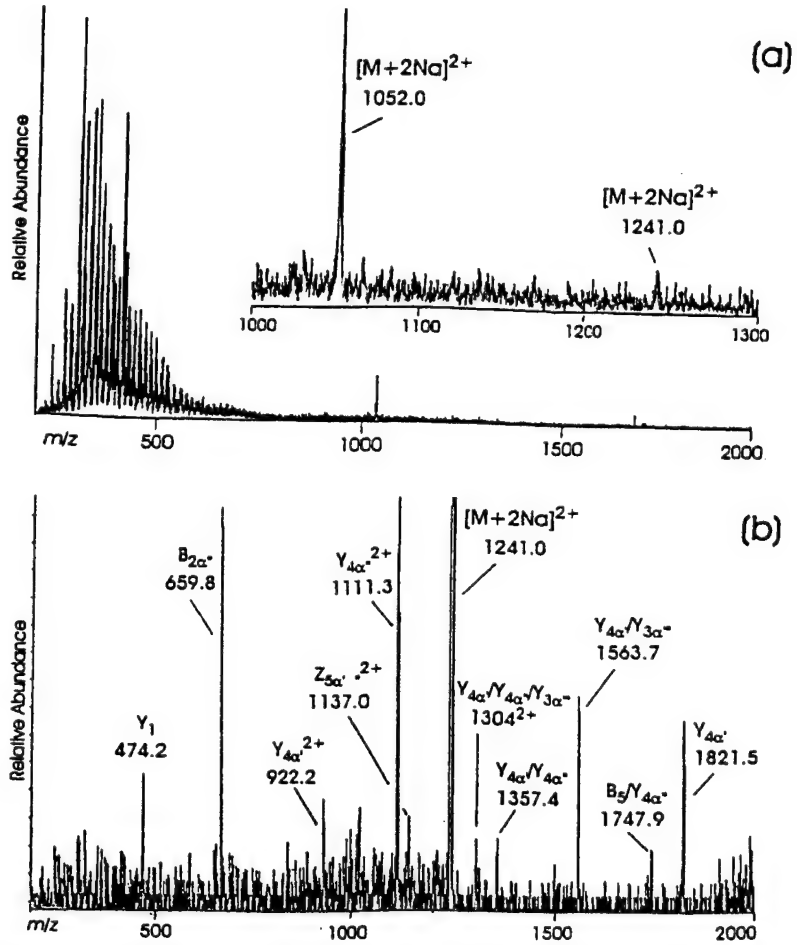
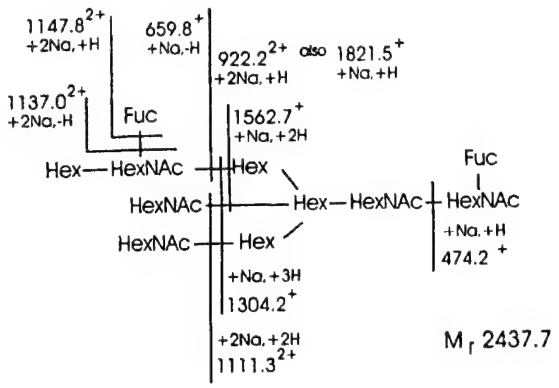


Fig. 3. (a) ESI-MS of the oligosaccharides released from a rat pineal gland glycopeptide and permethylated [33]. (b) CID mass spectrum of the $[M + 2Na]^{2+}$ ion of the minor component, m/z 1241.0. Product ion assignments are shown in Scheme 2Scheme 3.

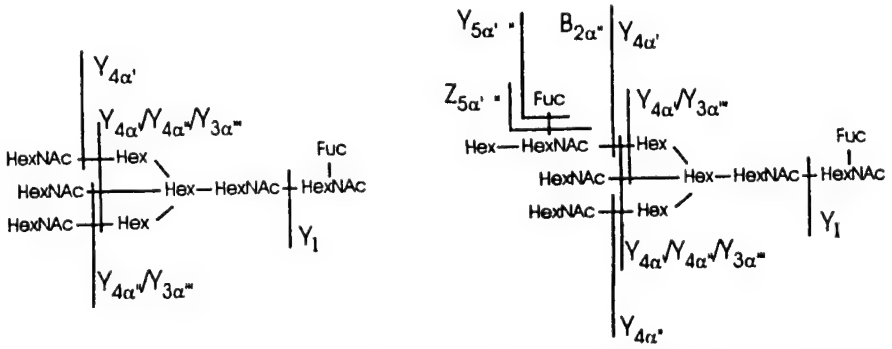


Scheme 2. Assignments of the fragment ions in the ESI-CID MS/MS spectrum of the permethylated minor glycoform released from the CPP glycopeptide (Fig. 5).

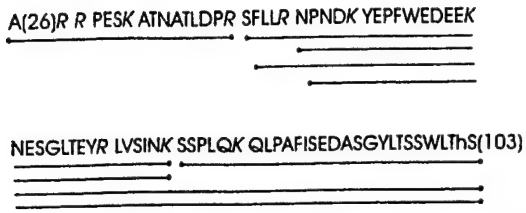
ciation (CID). Even for the minor oligosaccharide, the CID mass spectrum in Fig. 3(b) had excellent signal-to-noise ratio. The principal fragmentation pathway, glycosidic cleavage, yielded singly- and doubly-charged product ions that retained sodium (Scheme 2). The most significant fragments observed in the MS/MS spectra of the two glycoforms are indicated in Scheme 3, using the nomenclature system of Domon and Costello [35]; formation of *Y*-type ions involves cleavage of the glycosidic bond on the non-reducing-end side of the oxygen and includes hydrogen transfer to the observed fragment, whereas formation of *Z*-type fragments involves cleavage on the reducing-end side of the oxygen accompanied by hydrogen loss. *B*-type fragmentation does not involve hydrogen transfer. The spectra thus served to determine the residue sequences and branching patterns (but not the linkage sites). When adequate material is available to permit a second set of experiments, opening of the carbohydrate rings by periodate treatment can be used to provide linkage information [34,36], but the technique was not used in this case, because initial uncertainty about the amount and complexity of the sample suggested a conservative approach.

3.4. Cleavage of the activated thrombin receptor

The thrombin receptor (hTR) is a membrane-bound glycoprotein, with an exposed N-terminal region and seven transmembrane segments [37]. Binding of thrombin leads to cleavage of the R41–S42 bond and may lead to formation of an extracellular loop by implantation of the newly exposed terminus into the membrane [38,39]. While the receptor remains in this activated form, the membrane-bound G-protein exercises its signalling function and numerous physiological responses are stimulated, including blood coagulation, inflammation and mitosis. The process continues until the receptor is deactivated by an unknown mechanism. Because prolonged activation of the receptor can have deleterious effects, it would be advantageous to understand as completely as possible the mechanisms for activation and deactivation of the receptor. Our collaborator, A. Kuliopoulos at Tufts-New England Medical Center, is exploring the hypothesis that deactivation occurs upon further protease cleavage within the receptor's exodomain. As part of the investigation,



Scheme 3. Carbohydrate sequence and branching patterns assigned on the basis of fragment ions observed in the ESI-CID MS/MS spectra of the major (upper) and minor (lower) oligosaccharide glycoforms released from the CPP glycopeptide from rat pineal gland and permethylated [33].



Scheme 4. Amino acid sequence of the thrombin receptor exodomain (expressed in *E. coli* and cleaved from the carrier by CNBr), showing potential α-thrombin cleavage sites in italics. Single-letter amino acid codes are given; hS = homoserine. Peptides observed by MALDI-TOF MS are underlined [40].

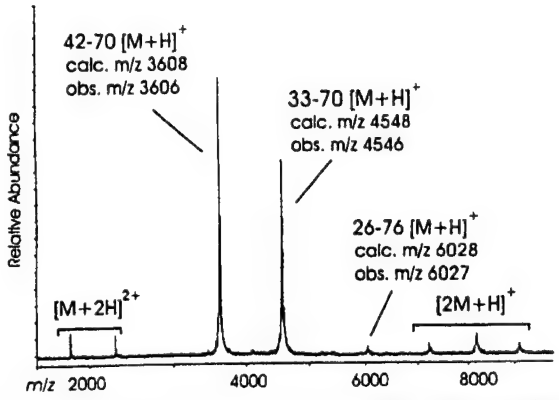


Fig. 4. MALDI-TOF mass spectrum of an HPLC fraction obtained after plasmin cleavage of the segment corresponding to the activated thrombin receptor exodomain (residues 42–103) [40].

the products of proteolysis with the various proteases known to be present, including thrombin itself, are being determined by MALDI-TOF MS of fractions separated by HPLC. The kinetics governing the appearance and disappearance of the various peptide products are being followed by HPLC. Because the sequence of the starting protein (corresponding to the external portion of the receptor, residues 26–103, a segment that has been expressed in *E. coli*) is known (Scheme 4), the molecular weights of potential products from cleavage with each protease can be calculated. In Scheme 4, spacings illustrate the peptides that would result from cleavage of the receptor exodomain at all the potential α-thrombin cleavage sites, the italicized *R* and *K* residues. MALDI-TOF MS analysis of the proteolysis products can identify the peptide(s) present in each of the HPLC fractions. This provides information on which of the potential cleavages occur under the experimental conditions and facilitates the kinetic studies. The underlines with circles at their termini in Scheme 4 indicate the peptides experimentally observed in α-thrombin digests, showing that not all the cleavage sites are actually utilized [40].

The MALDI-TOF mass spectra are simple, dominated by $[M_x + H]^+$ ions corresponding to each of the components in the HPLC fraction. Additional low abundance peaks can be assigned to doubly-charged species $[M_x + 2H]^{2+}$ and dimers $[2M_x + H]^+$, which may include mixed pairs of the various species $[M_x + M_y + H]^+$. As an example, Fig. 4 shows the MALDI-TOF mass spectrum of an HPLC fraction obtained after plasmin cleavage of the segment corresponding to the activated receptor exodomain (residues 42–103). Major components in these fractions are present at the pmol level; since the useable dynamic range is large, minor components may easily be detected at low fmol levels. The mass spectra thus provide sensitive and specific information on protease digestion products and permit the use of HPLC for the necessary kinetic analyses [40].

3.5. *Pegylation of proteins*

In addition to the requirement to determine naturally-occurring posttranslational modification of proteins, the analyst may also be faced with the problem of determining the extent of deliberate chemical modifications undertaken to change the circulation half-life of a protein or its biophysical properties. One such modification is the substitution of the polymer polyethylene glycol (PEG) onto the amino group of lysine residues. When the protein has use as a pharmaceutical, this modification may prolong its circulation, lower the dosage amount and frequency and facilitate the maintenance of stable therapeutic levels. The modification also changes solubility properties, and permits studies and utilization of enzymes in organic solvents. Nonaqueous media can be useful to probe biochemically significant enzyme intermediates. In both cases, for the sake of reproducibility and data interpretation, it is important to ascertain the level of incorporation of the modifier. MALDI-TOF MS has been shown to be useful for

this purpose [41,42].

Choice of measurement conditions (solvent, sample concentration, matrix, wavelength, laser power, analyzer) can influence results with this method; the selection becomes especially critical for the analysis of complex mixtures. Fig. 5(a) shows the ultraviolet (337 nm) MALDI-TOF mass spectrum obtained in the reflectron mode with 'super DHB' matrix (a 9:1 mixture of 2,5-dihydroxybenzoic acid and 2-hydroxy-5-methoxybenzoic acid [44]) for a sample of horseradish peroxidase modified under conditions selected to introduce only a few moles of PEG per mole of protein [41,43]. Peak width is limited by the half-width of the distribution of the PEG monomers. The distribution of the different species, spaced at intervals of 5000 daltons, the average molecular weight of the PEG, can easily be determined. This spectrum compares favorably with an earlier spectrum, Fig. 5(b), that we provided to this collaborator. The spectrum in Fig. 5(b) had been obtained for a similar sample using sinapinic acid matrix and a Vestec VT2000 linear MALDI-TOF MS, also operated at 337 nm [41].

3.6. Distribution of glycoprotein oligosaccharides

The extent of glycosylation and the glycoform distribution on a glycoprotein vary with the source of the glycoprotein and with the state of the organism or individual at the time the protein was processed—the patterns are affected, for example, by age, hormonal and immunological stresses, genetic mutations and specific tissue variations. By no means are the mechanisms nor all the effects of glycosylation changes yet fully understood. For structure–activity relationships to be established with a degree of confidence and pattern changes to be monitored, rapid and reliable methods for the determination of glycosylation patterns are required. In addition, since glycosylation can affect glycoprotein tissue distribution and half-life, the pattern expressed on recombinant proteins produced in batch processes for therapeutic uses should be determined initially for each protocol and monitored over the course of time. A simple approach for determination of carbohydrate moieties present on a glycoprotein (or other glycoconjugate) is the chemical or enzymatic release of the total oligosaccharide pool, followed by analysis of the released carbohydrates. Fig. 6 shows the ESI and MALDI mass spectra obtained for one such recombinant glycoprotein sample, after *N*-glycanase release of the oligosaccharides, followed by permethylation [45]. Once one includes a calculation that takes into account the distribution of components in the ESI spectrum over multiple charge states, the results of the two methods are in fact quite consistent. In these spectra, *N*-acetyl neuraminic acid (NA), known to be a labile group [46], was retained, protected as its fully methylated derivative.

In order to evaluate whether these methods give

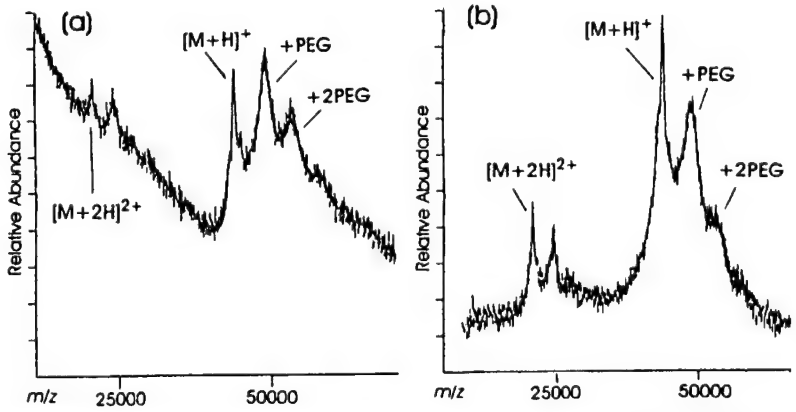


Fig. 5. MALDI-TOF mass spectra of different batches of horseradish peroxidase, M_r = 43,400, modified with polyethylene glycol, M_w = 5000, [41] dissolved in 0.5% trifluoroacetic acid recorded (a) for 3 pmol, in the linear mode with sinapinic acid matrix [41] and (b) for 200 fmol, in the reflector mode with 'super DHB' matrix [43].

abundance results equivalent to the widely-used fluorophore-assisted carbohydrate electrophoresis method (FACE), we have analyzed a mixture of glucose oligomers ($n = 2 \rightarrow 10$) as their permethylated derivatives [45]. In Fig. 7, the relative abundances observed for the signals attributed to the several oligomers are plotted for each of the methods. The data indicate some variance for the smallest components that probably results from their volatility. At the high mass end of the distribution, the initial experiment (5 keV accelerating energy, 5 kV postacceleration before the detector) suggested discrimination in the MALDI-TOF mass spectrum. When the accelerating voltage was increased from 5 to 10 kV (or the postacceleration was raised), however, this phenomenon was eliminated, indicating that the overall energy of the ions impacting on the detector is important to maintain the sensitivity for larger species.

Compared to other types of mass spectral analysis, ESI and MALDI are less prone to competitive effects in the ionization process, but some discrimination does still occur and its extent must be evaluated before quantitative measurements can be made. Carbohydrates bearing negatively-charged substituents (e.g., sulfate, phosphate) [47] or residues (e.g., *N*-acetyl neuraminic acid, glucuronic acid) [46] are more efficiently observed in the negative-ion mode.

For glycoproteins with multiple glycosylation sites, a more complete approach involves protease or chemical cleavage of the peptide backbone as a first step, because the glycosylation pattern is usually site-specific and information is lost by releasing the entire oligosaccharide pool at once. HPLC can provide separation of the individual glycosylation sites, and lectin affinity columns may be useful to separate glycotypes (e.g., biantennary vs. multiantennary, complex vs. high mannose). The molecular weight distributions of the eluted fractions may be determined for the glycopeptides [48], or for oligosaccharides released from them. Details of the structures may then be established using MS/MS [34], post-source decay [7] and/or additional measurements to determine mass shifts after employing a series of exoglycosidases [49–51].

3.7. Muramyl peptides in wild and mutant bacterial spores

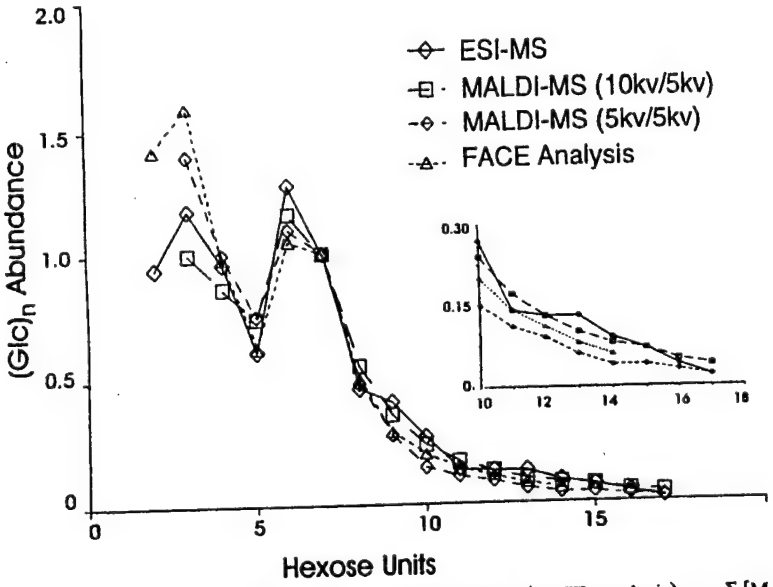


Fig. 7. Comparison of signal intensities observed for the fluorescence signal (FACE analysis) or $\Sigma_z[M + zNa]^{z+}$ of ions (MS) corresponding to components ($n = 2 \rightarrow 17$) for a permethylated $(glucose)_n$ polymer sample, each method's curve normalized against its value for $n = 7$.

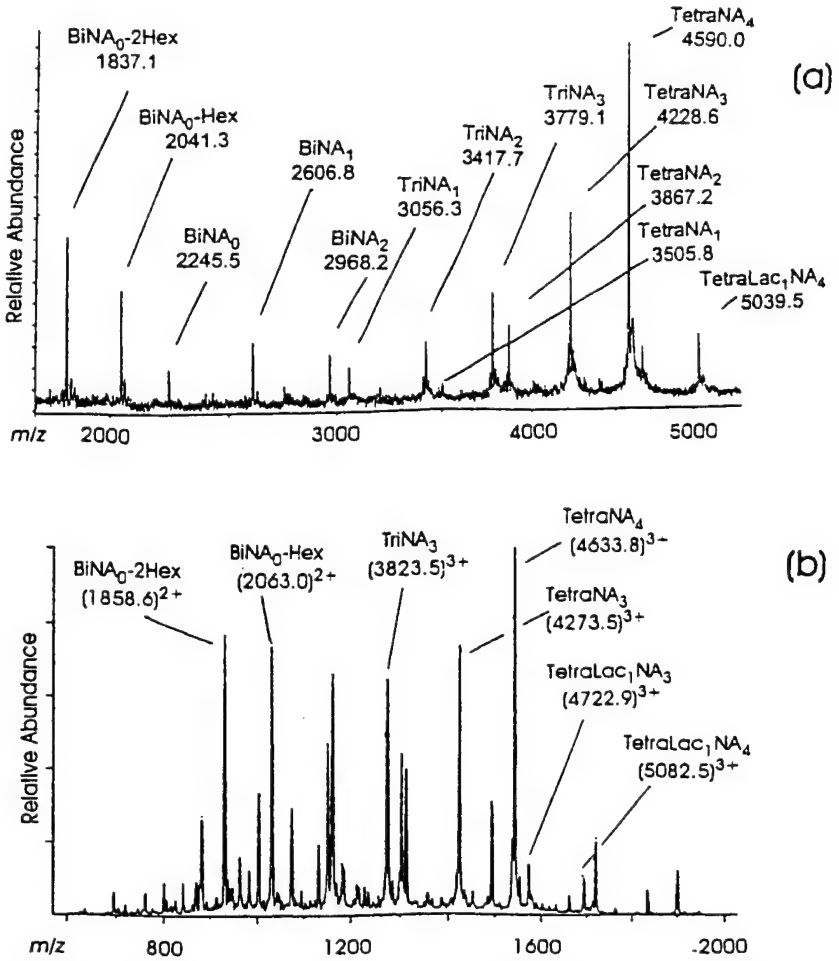
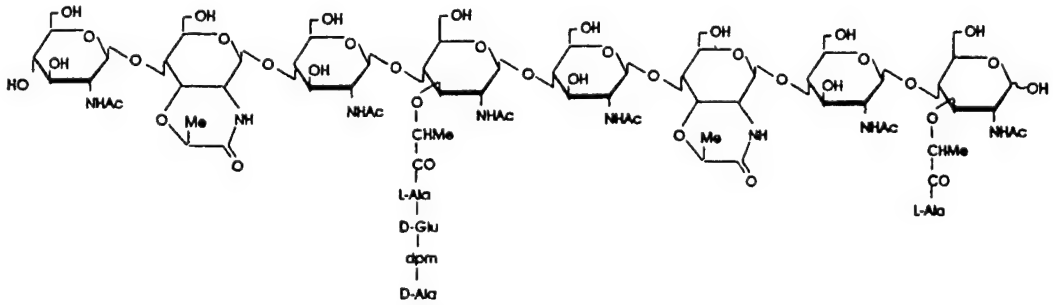


Fig. 6. Mass spectra of the oligosaccharide pool permethylated after *N*-glycanase release from a recombinant protein, LFA-3. Bi = biantennary; Tri = triantennary; Tetra = tetraantennary; NA = *N*-acetyl neuraminic acid; Lac = *N*-acetyllactosamine. Experimental values are given for mass assignments. (a) MALDI-TOF MS. All assignments correspond to $[M + Na]^+$. (b) ESI-MS. All assignments correspond to $[M + zNa]^{z+}$. Because of the space limitation within the figure, only the most abundant charge state has been marked for each component; undesignated peaks correspond to additional charge states of the various glycoforms. In order to obtain the relative abundance in the mixture for each component, the signals for all of its charge states (e.g., $z = 1–4$) must be summed.



Scheme 5. General structure for the muramyl peptides from bacterial spore peptidoglycan, proposed by Warth and Strominger [52].

Gram-positive bacteria, e.g., *Clostridium*, *Bacillus*, form endospores as a means of shielding themselves from environmental challenges during a dormant state. The survival and long-term viability of the species may depend on the structure of the spore, and, in particular, on the peptidoglycan (PG) that makes up most of the bacterial cell's thick outer protective layer, the cortex [52,53]. Our collaborators, D. Popham and P. Setlow at the Univ. of Connecticut Health Center, have prepared cultures of wild and mutant bacteria in order to correlate variations in the endospore peptidoglycan with survival. The carbohydrate portion of the glycopeptide is composed of alternating muramic acid and *N*-acetyl glucosamine residues (Scheme 5); the muramic acid residues may be modified with amino acids or short peptides that can crosslink. Unsubstituted muramic acid residues can cyclize to lactams. In order to probe details of the cortex structure, it has been necessary to devise sensitive and specific methods to determine the nature and frequency of these crosslinks. Enzymatic digestion of the PG with muramidase, followed by HPLC separation of the released glycopeptides (after borohydride reduction to remove the anomeric heterogeneity) yields profiles of the PG components. MALDI-TOF MS analysis of the HPLC fractions provides molecular weight information that can be interpreted in terms of carbohydrate and amino acid composition, and reveals the presence of lactams and their reduction products. Because acidic residues (glutamic acid, diaminopimelic acid) occur frequently in these glycopeptides, they can efficiently generate stable carboxylate anions. Negative-ion MALDI-TOF analysis is therefore particularly useful for highly sensitive molecular weight determinations that are little affected by residual contaminants.

Previous reports had demonstrated that the structures of muramyl peptides may be determined at the nmol level by liquid secondary ionization [LSI] tandem mass spectrometry [54]. Our recent experiments have shown that MALDI-PSD-TOF MS yields similar information on sequence and cross-linking, but at the pmol level [55]. The PSD spectrum of a PG tetrasaccharide tetrapeptide from wild-type *Bacillus* spore cortex is given in Fig. 8. Because precursor ion selection was not employed in this case, fragments from both the $[M+H]^+$ ion, m/z 1362, and the

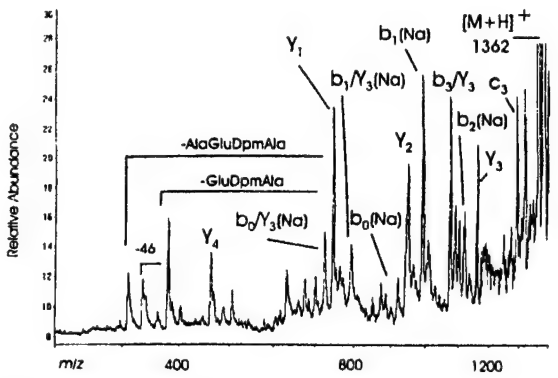
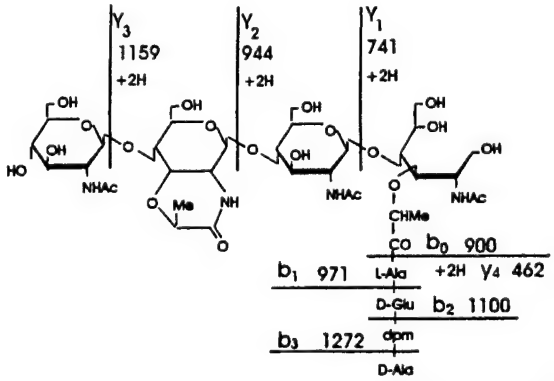


Fig. 8. MALDI-PSD-TOF mass spectrum of the $[M+H]^+$, m/z 1362, from a peptidoglycan tetrasaccharide tetrapeptide from a muramidase digest of wild-type *B. subtilis* spore cortex. Some contributions from the $[M+Na]^+$, m/z 1384, are also present. Peak assignments are shown in Scheme 6. Nitrogen laser, 337 nm. Accelerating voltage 10 kV.

$[M+Na]^+$ ion, m/z 1384 were observed. Most fragments originated from the less stable protonated species, as expected. Fragmentation in both the oligosaccharide and peptide portions of the structure permits the entire sequence to be deduced, and presence of the lactam to be observed, as indicated in Scheme 6. In this scheme, upper-case letters designate carbohydrate fragmentations [35] and lower-case letters designate peptide bond cleavages [56]. This sensitivity makes feasible determinations that would otherwise have required preparation and extraction of prohibitive amounts of the bacterial cultures.

These investigations have verified some of the



Scheme 6. Assignments of product ions in the MALDI-PSD-TOF mass spectrum (Fig. 9) of a tetrasaccharide tetrapeptide from peptidoglycan of wild-type *Bacillus* [55].

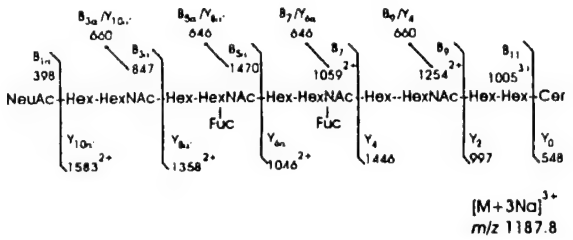
assumptions made in previous work and provide tools for wider studies [53]. One recent phase of this continuing collaboration showed that a long-held assumption is incorrect, when we determined that cortex PG cross-linking is necessary but does not control the rate of dehydration in the dormant state, and is instead critically related to the success of reproduction following a dormant period [57].

3.8. Studies of glycosphingolipids

Glycosphingolipids (GSLs) play many roles, from nervous system signalling to intercellular communications [58]. Heterogeneity is present in both the carbohydrate and lipid portions, and varies with tissue source, age, and immune status. Molecular weight profiles of such materials can be obtained by either MALDI-TOF MS [59,60] or ESI-MS [61]. Structural details must then be determined from fragments arising by CID or PSD.

Unusual structures are encountered in some instances, such as cell-surface GSLs on tumor tissues. For these, widely-used methods that base identifications on chromatographic behavior alone or in combination with exoglycosidase digestions are insufficient to specify unusual sequences and branching patterns. In a collaborative study carried out jointly with S.-I. Hakomori et al., at the Biomembrane Institute, Seattle, WA, human myelogeneous leukemia HL60 cells that bind to E-selectin and possibly P-selectin were shown to contain gangliosides with extended *N*-acetyllactosamine structures and varied fucosylations that were absent in normal leukocytes [62,63]. These represent a new type of epitope for E-selectin binding, different from the well-known sialylated Lewis^x ligand, Neu5NAcα2 $\rightarrow$ 3Galβ1 $\rightarrow$ 4(Fucα1 $\rightarrow$ 3)GlcNAcβ1 $\rightarrow$ 3R. The structures were elucidated based on ESI-MS/MS of major and minor GSLs, together with fast atom bombardment MS and NMR analysis of the more abundant components. The structure shown in Scheme 7 is representative of the set.

For an ongoing study that aims at delineating fucosyltransferase substrates in the developing brain, conducted in collaboration with R.H. McCluer et al. at the Shriver Center for Mental Retardation, Waltham, MA, it has been necessary to optimize methods for the analysis of GSLs at the low pmol

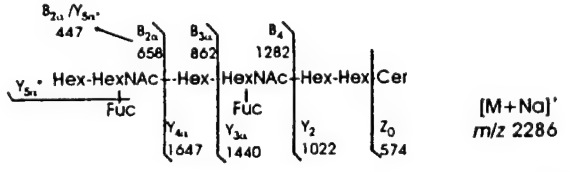


Scheme 7. Structure and CID product ion mass assignments for the permethylated derivative of a ganglioside from human myelogeneous leukemia HL60 cells that reacts with E-selectin [63].

level. To achieve sample cleanup and maximize sensitivity, the samples have been subjected to peracetylation, followed by permethylation. Molecular weight determinations have been made by LSI-MS on a four-sector tandem mass spectrometer and MALDI-TOF MS. Although LSI-MS/MS can be used for structural determinations of the more abundant model compounds, the sample amounts from foetal calf brain were too low for this method. Our model studies showed that MALDI-PSD-TOF MS can be useful for this class of compounds. PSD was performed on a sample corresponding to less than 700 fmol of one of the unknown calf brain GSLs and thereby enabled elucidation of its carbohydrate moiety (Scheme 8) [64]. PSD thus represents a valuable new tool for high sensitivity glycolipid structural determinations.

3.9. Emerging approaches for biopolymer analysis: Infrared-MALDI-TOF MS and FT-ICR-MS

MALDI-TOF mass spectra of ions desorbed with an infrared laser exhibit abundant multiply-charged species and multimers [65]. Decrease in metastable decomposition leads to improved peak shapes compared to UV-MALDI, especially noticeable for large



Scheme 8. Structure and PSD product ion mass assignments for the permethylated derivative of a neutral glycosphingolipid from foetal calf brain [64].

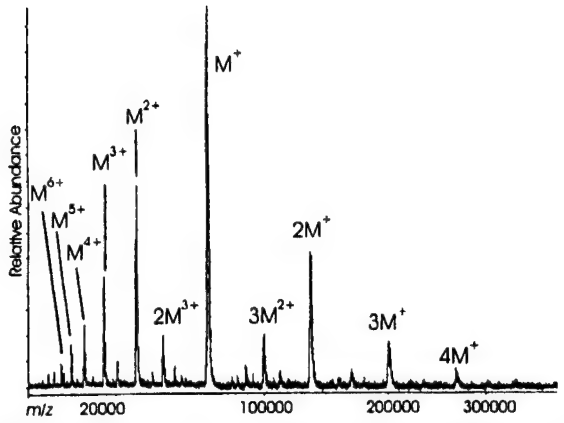


Fig. 9. MALDI-TOF mass spectrum of 1 pmol bovine serum albumin, M_r = 66,431, recorded with an infrared laser (2.94 μm) in the reflector mode, with succinic acid matrix.

molecules. As an example, Fig. 9 shows the 2.94 μm IR-MALDI mass spectrum of bovine serum albumin, recorded in the reflector mode, with succinic acid matrix. Infrared irradiation has been demonstrated to yield superior results for analysis of proteins and glycoproteins electroblotted onto transfer membranes, perhaps because of the deeper penetration of the infrared beam into the target surface [66].

The mass range for the TOF analyzer is theoretically unlimited, but is bounded by sample behavior and detector response. The practical upper mass limit for MALDI-TOF MS is around 1×10^6 daltons, a figure that has been achieved for IgM [67], but most bipolymer measurements to date have been in the range below 400 kDa. For industrial types of polymers, MALDI-TOF MS measurements of high mass compounds, even exceeding 1 MDa have been reported, although the dominant species for the very large ions bore two or more charges [68,69]. The mass range with other types of analyzers is more limited; for example, the highest singly-charged biopolymer ions observed by FT-ICR MS thus far are oligonucleotide 25-mers at m/z 7634 [70]. Multiple-charging in ESI-FT-ICR mass spectra extends the effective mass range, so that spectra of two different porcine serum albumins, M_r 66,736 and 66,886, have been obtained with isotopic resolution [71]. These instruments offer the possibility for multiple stages of mass analysis as well as the advantage of very high resolution and mass accuracy. Masses that are measured with the high accuracy provided by FT-ICR MS, sector instruments, or TOF MS with delayed extraction yield information on elemental composition in the lower mass ranges, and narrow the database search windows for larger compounds such as proteins [72].

4. Concluding remarks

Powerful new mass spectral ionization techniques, in particular electrospray and matrix-assisted laser desorption, together with microscale methods for sample preparation and improvements in mass analyzers, can have impact on wide areas of biological science. This brief survey of a few of the projects in which our own laboratory has recently been engaged is meant to provide a window on the types of 'real-world' problems that can now be approached. Other important areas include characterization of oligonucleotides and of non-covalent complexes. Developments are expected to continue at a rapid pace worldwide, increasing even further the reach of mass spectrometry into biology and medicine.

Acknowledgements

The author wishes to thank her colleagues J. Helin, X. Hronowski, P. Juhasz, L.C.N. Ngoka, H. Perreault, B.B. Reinhold, V.N. Reinhold and S. Ye, and the collaborators named in the text, for their contributions to the work described. The research was supported by the National Institutes of Health Grants P41 RR10888 and S10 RR10493 (to C.E. Costello), P41 RR00317 (to K. Biemann); R29 GM52926 (to A. Kuliopoulos) and R01 CA42505 (to S.-I. Hakomori); NATO Collaborative Grant 900605 (to C.E. Costello); National Science Foundation Grant No. MCB 9400633 (to V.N. Reinhold); the Vrie Univ., Amsterdam; the Swedish Medical Research Council (to R. Ekman) and Thermo BioAnalysis J. Helin and H. Perreault received fellowships from the Academy of Finland and the National Research Council of Canada, respectively.

References

[1] M. Karas, D. Bachmann, U. Bahr, F. Hillenkamp, Int. J. Mass Spectrom. Ion Proc. 78 (1987) 53.
[2] J.B. Fenn, M. Mann, C.K. Meng, S.F. Wong, C.M. Whitehouse, Science 246 (1989) 64.
[3] A. Varki, Glycobiology 3 (1993) 97.
[4] K.G. Standing, R.C. Beavis, G. Bollbach, W. Ens, F. Lafortune, D. Main, B. Schueler, X. Tang, J.B. Westmore, Anal. Instrumen. 16 (1987) 173.
[5] X. Tang, W. Ens, K.G. Standing, J.B. Westmore, Anal. Chem. 60 (1988) 1791.
[6] R. Kaufmann, B. Spengler, F. Lützenkirchen, Rapid Commun. Mass Spectrom. 7 (1993) 902.
[7] B. Spengler, D. Kirsch, R. Kaufmann, J. Lemoine, J. Mass Spectrom. 30 (1995) 782.
[8] W.C. Wiley, I.H. McLaren, Rev. Sci. Instr. 26 (1953) 1150.
[9] R.M. Whittal, L. Li, Anal. Chem. 67 (1995) 1950.
[10] R.S. Brown, J.J. Lennon, Anal. Chem. 67 (1995) 1998.
[11] M. Wilm, A. Shevchenko, T. Houthaeve, S. Breit, L. Schweigerer, T. Fostis, M. Mann, Nature 379 (1996) 466.
[12] K.J. Light-Wahl, B.L. Schwartz, R.D. Smith, J. Am. Chem. Soc. 116 (1994) 5271.
[13] N.L. Kelleher, C.A. Costello, T.P. Begley, F.W. McLafferty, J. Am. Soc. Mass Spectrom. 6 (1995) 981.
[14] H. Görisch, Anal. Biochem. 173 (1988) 393.
[15] J.C. Rouse, J.E. Vath, Anal. Biochem. 238 (1996) 82.
[16] E. Nordhoff, A. Ingendoh, R. Cramer, A. Overberg, B. Stahl, M. Karas, F. Hillenkamp, P.F. Crain, Rapid Commun. Mass Spectrom. 6 (1992) 771.
[17] I. Ciucanu, F. Kerek, Carbohydr. Res. 131 (1984) 209.
[18] K.B. Linsley, S.-Y. Chan, S. Chan, B.B. Reinhold, P.J. Lisi, V.N. Reinhold, Anal. Biochem. 219 (1994) 207.
[19] H. Perreault, C.E. Costello, Org. Mass Spectrom. 29 (1994) 720.
[20] D. Liu, C. Dass, G. Wood, D.M. Desiderio, J. Chromatogr. 500 (1990) 395.
[21] R. Ekman, Nord. Psykiatr. Tidskr. 39 (1985) 31.
[22] J.Z. Chou, M.J. Kreek, B.T. Chait, J. Am. Soc. Mass Spectrom. 5 (1994) 10.
[23] C.R. Jiménez, P.A. van Veelen, K.W. Li, W.C. Wildering, W.P.M. Geraerts, U.R. Tjaden, J. van der Greef, J. Neurochem. 62 (1994) 404.
[24] P. Adrén, R.M. Caprioli, J. Mass Spectrom. 30 (1995) 817.
[25] R. Ekman, P. Juhasz, M. Heili, H. Ågren, C.E. Costello, Peptides 17 (1996) 1107.
[26] K. Tatemoto, M. Carlquist, V. Mutt, Nature 296 (1982) 659.
[27] S.H.C. Hendry, in: W.F. Colmers, C. Wahlestedt (Eds.), The Biology of Neuropeptide Y. Humana Press, Totowa, NJ, 1993, p. 65.
[28] B.G. Stanley, in: W.F. Colmers, C. Wahlestedt (Eds.), The Biology of Neuropeptide Y. Humana Press, Totowa, NJ, 1993, p. 457.
[29] M. Heilig, G.F. Koob, R. Ekman, K.T. Britton, TINS 17 (1994) 80.
[30] E. Widerlöv, L.H. Lindström, C. Wahlestedt, R. Ekman, J. Psychiat. Res. 22 (1988) 69.
[31] W.H. Berrettini, A.R. Doran, J. Kelsoe, A. Roy, D. Pickar, Neuropsychopharmacology 1 (1987) 81.
[32] L. Grundemar, S.P. Sheikh, C. Wahlestedt, in: W.F. Colmers, C. Wahlestedt (Eds.), The Biology of Neuropeptide Y. Humana Press, Totowa, NJ, 1993, p. 197.
[33] C.R. Jiménez, K. Dreisewerd, H.D. Mansvelder, A.B. Brussard, B.B. Reinhold, C.E. Costello, R. Kingston, M. Karas, F. Hillenkamp, J.P.H. Burbach, K.W. Li, W.P.M. Geraerts, Abstr. Desorption 96, Bornholm, Denmark, September, 1996.
[34] V.N. Reinhold, B.B. Reinhold, C.E. Costello, Anal. Chem. 67 (1995) 1772.
[35] B. Domon, C.E. Costello, Glycoconj. J. 5 (1988) 397.
[36] A.S. Angel, B. Nilsson, Biomed. Environ. Mass Spectrom. 19 (1990) 721.
[37] T.-K. Vu, D.T. Hung, V.I. Wheaton, S.R. Coughlin, Cell 64 (1994) 1057.
[38] R.E. Gerszten, J. Chen, M. Ishii, K. Ishii, L. Wang, T. Nanevicz, C.W. Turk, T.-K.H. Vu, S.R. Coughlin, Nature 368 (1994) 648.
[39] T. Nanevicz, M. Ishii, L. Wang, M. Chen, C.W. Turk, F.E. Cohen, S.R. Coughlin, J. Biol. Chem. 270 (1995) 21619.
[40] A. Kuliopulos, P.J. Sheridan, S.K. Seeley, J. Helin, C.E. Costello, Abstr. Am. Soc. Hematology, November, 1996.
[41] P.A. Mabrouk, J. Am. Chem. Soc. 117 (1995) 2141.
[42] S.K. Chowdhury, M. Doleman, D. Johnston, J. Am. Soc. Mass Spectrom. 6 (1995) 478.
[43] C.E. Costello, J. Helin, L.C.M. Ngoka, Progr. Biomed. Opt. 2680 (1996) 406.
[44] M. Karas, H. Ehring, E. Nordhoff, B. Stahl, K. Strupat, F. Hillenkamp, M. Grehl, B. Krebs, Org. Mass Spectrom. 28 (1993) 1476.
[45] B.B. Reinhold, S. Ye, J. Helin, V.N. Reinhold, C.E. Costello, Proc. ASMS Conf. on Mass Spectrom. Allied Topics 44 (1996) 222.
[46] M.C. Huberty, J.E. Vath, W. Yu, S.A. Martin, Anal. Chem. 65 (1993) 2791.
[47] G.R. Hayes, A. Williams, C.E. Costello, C.A. Enns, J.J. Lucas, Glycobiology 5 (1994) 227.
[48] M.J. Treuheit, C.E. Costello, H.B. Halsall, Biochem. J. 283 (1992) 105.
[49] B. Stahl, T. Klabunde, H. Witzel, B. Krebs, M. Steup, M. Karas, F. Hillenkamp, Eur. J. Biochem. 220 (1994) 321.
[50] E. Mørtz, T. Sareneva, I. Julkunen, P. Roepstorff, J. Mass Spectrom. 31 (1996) 1109.
[51] R. Küster, T.J.P. Naven, D.J. Harvey, J. Mass Spectrom. (1996) 1131.
[52] A.D. Warth, J.L. Strominger, Biochemistry 11 (1972) 1389.
[53] D.L. Popham, J. Helin, C.E. Costello, P. Setlow, J. Bacteriol. 178 (1996) 6451.
[54] S.A. Martin, R.S. Rosenthal, K. Biemann, J. Biol. Chem. 262 (1987) 7514.
[55] J. Helin, C.E. Costello, D. Popham, P. Setlow, Proc. ASMS Conf. on Mass Spectrom. Allied Topics 44 (1996) 1000.
[56] K. Biemann, in: J.A. McCloskey (Ed.), Mass Spectrometry (Methods in Enzymology, Vol. 193) Academic Press, San Diego, 1990, p. 886.
[57] D.L. Popham, J. Helin, C.E. Costello, P. Setlow, Proc. Natl.

Acad. Sci. USA (1997) in press.
[58] S.-I. Hakomori, J. Biol. Chem. 265 (1990) 18713.
[59] H. Egge, J. Peter-Katalinic, M. Karas, B. Stahl, Pure Appl. Chem. 69 (1991) 491.
[60] P. Juhasz, C.E. Costello, J. Am. Soc. Mass Spectrom. 3 (1992) 785.
[61] B.B. Reinhold, S.-Y. Chan, S. Chan, V.N. Reinhold, Org. Mass Spectrom. 29 (1994) 736.
[62] M.R. Stroud, K. Handa, M.E.K. Salyan, K. Ito, S.B. Levery, S.-I. Hakomori, B.B. Reinhold, V.N. Reinhold, Biochemistry 35 (1996) 758.
[63] M.R. Stroud, K. Handa, M.E.K. Salyan, K. Ito, S.B. Levery, S.-I. Hakomori, B.B. Reinhold, V.N. Reinhold, Biochemistry 35 (1996) 770.
[64] H. Perreault, O. Koul, J. Street, R.H. McCluer, C.E. Costello, Proc. ASMS Conf. on Mass Spectrom. Allied Topics 43 (1995) 1179.
[65] A. Overberg, M. Karas, U. Bahr, R. Kaufmann, F. Hillenkamp, Rapid. Commun. Mass Spectrom. 4 (1990) 293.
[66] K. Strupat, C. Eckerskorn, M. Karas, F. Hillenkamp, in: A.L. Burlingame, S.A. Carr (Ed.), Mass Spectrometry in the Biological Sciences, Humana Press, Clifton, NJ, 1995, p. 365.
[67] R.W. Nelson, D. Dogruel, P. Williams, Rapid. Commun. Mass Spectrom. 9 (1995) 625.
[68] P.O. Danis, D.E. Carr, Org. Mass Spectrom. 28 (1993) 923.
[69] D.C. Schriemer, L. Li, Anal. Chem. 68 (1996) 2721.
[70] Y. Li, K. Tang, D.P. Little, H. Köster, R.L. Hunter, R.T. McIver Jr., Anal. Chem. 68 (1996) 2090.
[71] F.W. McLafferty, J. Mass Spectrom. 30 (1995) 39.
[72] O.N. Jensen, A. Podtelejnikov, M. Mann, Rapid Commun. Mass Spectrom. 10 (1996) 1371.

Blank

NEXT-GENERATION BIOLOGICAL MASS SPECTROMETER

S. Randolph Long
Edgewood RDE Center
Aberdeen Proving Ground, MD 21010

ABSTRACT

Enhancements in the defensive posture against agents of biological origin that should accrue from the development of a next-generation biological mass spectrometer are described. Overall features expected to be incorporated in such an instrument are briefly outlined along with identification of key issues in the development of this type of instrument.

I. INTRODUCTION: THE CHEMICAL BIOLOGICAL BATTLEFIELD AND THE ROLE OF MASS SPECTROMETRY

The challenge of rapidly detecting and identifying a potential biological agent aerosol with sufficient time and accuracy to permit service personnel to assume an adequate protective posture is one of great complexity. The widely held view is that this challenge must be addressed through the application of a network of sensors providing both an early warning capability to alert personnel as soon as possible to the potential of a wind-borne biological threat and a confirmation of the nature of the biological threat and its identification.

Approaches providing an early warning capability include "standoff" sensors, such as LIDAR using either near-infrared or ultraviolet laser radiation (providing a capability, respectively, of several tens of kilometers or several kilometers advanced warning of particulate material in the atmosphere), and "remoted" sensors, such as a distributed network of particle counters.

Confirmation of the biological nature of an aerosol and identification of its constituent(s) requires sampling and analysis of the aerosol. No single component system is yet available to address the identification of a biological aerosol; it is considered that the most logical approach for the near term must utilize a suite of component systems each performing a necessary function to provide most accurate response within a given time frame. Such a suite is likely to include a particle trigger to initiate the sampling of the aerosol, a "detector(s)" which provides rapid (ca 1-2 minute) confirmation that the aerosol is biological (and perhaps provide a finer classification of the aerosol as spores, Gram negative bacteria, or the like), and an identifier(s), which provides specific identification of the biological aerosol, on the timescale of ten minutes.

Presently used identification approaches utilize antibody-antigen recognition as the

specific identification mechanism. On the horizon are gene probe approaches which permit the very sensitive recognition of nucleic acid (DNA or RNA) fragments of the genome of bacteria and viruses. Both antibody and gene probe based approaches are very specific, requiring a different assay for each potential biological threat material.

With the advent of ionization techniques which permit the mass analysis of intact large biological molecules, it is now possible to consider the development of a next-generation mass spectrometer with a capability for specific identification of biological threats. In the potential role of an identification technology, a next-generation mass spectrometer based on these ionization approaches has four characteristics that, to our knowledge, no other identification technique has. These characteristics include: (1) the capability to identify a wide range of agents of biological origin (protein toxins, bacteria, viruses) with potentially a single assay, (2) the capability to characterize unknown or unanticipated threats, (3) the capability to perform a field update of its library in response to previously unknown/uncharacterized threats, (4) with minor modifications, the addition of a capability to detect/identify chemical threats as well as biological threats. Thus, next-generation biological mass spectrometry occupies a special niche in the challenge of biological detection/identification.

II. CURRENT APPROACHES TO BIOLOGICAL MASS SPECTROMETRY IN THE FIELD

To use the term "next-generation" biological mass spectrometer implies the existence of a current generation biological mass spectrometer. Indeed, such a capability is currently in advanced development for inclusion in a U. S. Army fielded biological detection suite. This instrument, known as the Chemical Biological Mass Spectrometer (CBMS), utilizes pyrolysis to vaporize the components of an aerosol collected by impaction of the effluent of an aerosol sampler onto a quartz fiber plug. The stable pyrolysates are then passed through a transfer line into an ion trap mass spectrometer, where they are analyzed by electron impact ionization.

The CBMS is to be used as a "detector" subsystem in a fielded biodetection suite. As such, it will provide rapid classification of collected aerosol as biological or not; it can provide further information with regard to classification of aerosol generically as spores, cells, or toxin. Testing of the CBMS approach has demonstrated some capability to differentiate among major target cellular pathogens.

The rather long (millisecond timescale) infrared pyrolysis used to vaporize biological materials in the CBMS leads to the production of relatively low mass stable pyrolysates, with resultant mass spectra of biological materials comprising peaks limited to a maximum of about 250 Da with a large degree of spectral congestion and complexity. These masses are much lower than those associated with the major biomolecules, such as proteins and lipids, that are found in agents of biological origin (ABOs). The application of fatty acid methyl esterification in the CBMS has produced encouraging results toward enhancing bacterial differentiation based on fatty acid content. However, in order to mass analyze the high mass biomolecule constituents of ABOs without the significant degradation resulting from pyrolysis, different vaporization/ ionization techniques are required. With the advent of the electrospray ionization (ESI) and

matrix-assisted laser desorption ionization (MALDI) techniques, it has become possible to readily analyze mass spectrometrically intact large biomolecules. Techniques such as these new ionization approaches permit us to consider the conceptualization of a next-generation biological mass spectrometer with potential for identification of ABOs.

III. A NEXT-GENERATION BIOLOGICAL MASS SPECTROMETER

Other contributions to these Proceedings outline the ESI and MALDI techniques, so details will not be presented here. The specific advantage that these techniques have in vaporizing and ionizing large molecules without fragmentation is essentially that the processes of vaporization and ionization are coupled with mediation by a "matrix" which acts both to facilitate the ionization of the analyte and to remove excess energy associated with vaporization or ionization from the analyte. In ESI, the analyte is ionized by passing a dilute solution containing it in an appropriate solvent through a charged needle. The resultant spray comprises charged microdroplets which are subsequently desolvated, leaving bare analyte molecules charged in a distribution of charge states. In MALDI, the analyte is usually incorporated as a dilute solution in a solid matrix. Upon irradiation by a short low-energy laser pulse, a plume of desorbed matrix ions and analyte molecules is generated, followed rapidly by charge-transfer reactions leading to ionization of the analyte.

Application of either of these ionization approaches to agents of biological origin (ABOs: protein toxins, bacteria, viruses) requires that the biomolecular constituents of these materials be "extracted" and conveyed to the ionization source with appropriate steps to prepare them for mass spectrometric analysis. A diagrammatic presentation of a conceptual system to accomplish this is presented as Figure 1.

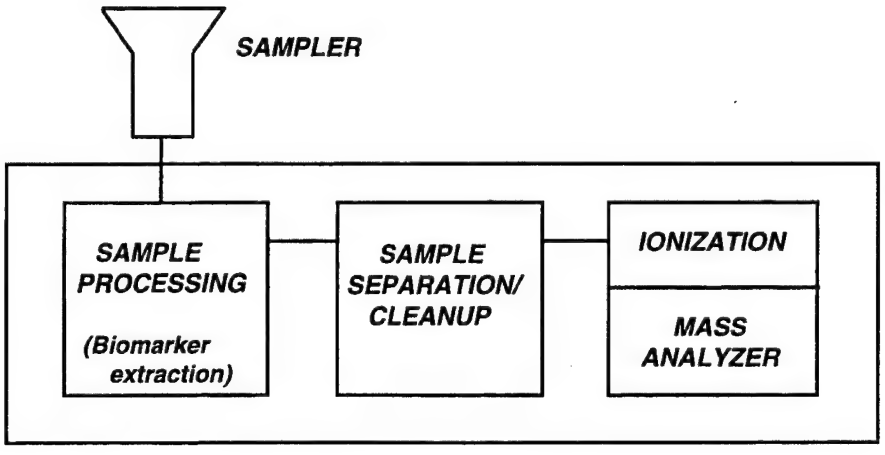


Figure 1

The air sampler depicted in Figure 1 may or may not be an integral part of the mass spectrometric system. The purpose of the sampler is to concentrate airborne particulates into a solvent system for further analysis. Current implementations in US Army biological detection suites utilize sampling systems which select the respirable air particles (1-10 micron diameter size range) and impact them into a small (ca 1 ml) volume of liquid, which then constitutes the sample to be analyzed for biological materials. Many variations on specific sampling methodologies are possible.

The remaining component "modules" represent conceptual processes which must be applied to analyze and classify/identify airborne particulates mass spectrometrically. These processes may occur in separate modules, as depicted, or in modules which combine some of the functions shown. The biomolecular constituents of the aerosol particles must first be extracted by some means from the particles in order to be presented to the mass spectrometer for analysis as intact separate entities. Prior to sample cleanup/preparation for mass analysis, it is very likely that a separation step must be performed since the mixture of analytes generated by the extraction process must be assumed to be rather complex, particularly for bacterial and viral aerosols. Sample cleanup must be performed to remove solution components that may hinder effective mass analysis. Typical biological solution components that have an impact on mass analysis include salts, detergents, and other materials that may be in the aerosol or may be used during the extraction and separation phases. Since the next-gen bio MS should be capable of largely unattended operation, it is important that all processes involved in sample processing and mass analysis be coupled in on-line fashion.

IV. KEY ISSUES IN THE DEVELOPMENT OF A NEXT-GENERATION BIOLOGICAL MASS SPECTROMETER

A number of key issues must be addressed in order to define optimum configuration(s) for a next-generation biological mass spectrometer. The Joint Services Workshop on Biological Mass Spectrometry was convened as a forum to present and discuss results in the biological mass spectrometry community which have an impact on these issues. Most of these are summarized in detail in these Proceedings. Here we enumerate and discuss briefly the major issues involved.

- Biomarker data base: Biomarkers comprise the ABO biomolecular constituents whose detection by mass spectrometry serves as the basis for differentiation/identification. Identification of these biomarkers is the most essential science base issue for next-generation mass spectrometry. Since the particular advantage that mass spectrometry has with regard to ABO identification is its potentially very broad range of application, the class of biomolecules selected as the principal target biomarker class should also have broad applicability. Ideally the biomarkers should be readily detected, have reasonably consistent abundance for the same ABO regardless of source and history, and be differentiable from those of ambient background materials as well as other ABOs. An additional important question related to biomarker database is the form of mass analysis required, i.e., whether mass spectral profiles are sufficient or tandem

mass spectrometry must be performed, and whether methods such as enzymatic digestion provide enhanced discrimination potential.

- Ionization approach: Two principal candidate approaches are available: ESI and MALDI. It is not feasible to use both in a next-generation bio mass spectrometer, hence, a choice must be made. An important consideration is that the next-gen bio MS must be capable of on-line coupling to sample processing. ESI therefore has a particular advantage in this respect, though MALDI provides a capability for simpler mixture analysis. It is possible that one approach is more valid for the short term, while further improvement in the second may make it more attractive in the future.

- Biomarker extraction approaches: Various approaches are available for each of the major biomarker classes (proteins, lipids, DNA). An emphasis must be placed on speed and simplicity of the chemistry involved, as well as on approaches that utilize minimal consumables and introduce fewest solution components that may have to be removed prior to mass analysis.

- Chemical separation: While a mass spectrometer is capable of a significant degree of mixture analysis itself, the effective analysis by a mass spectrometer alone of a large number of constituents as will be found in a bacterial sample is not feasible. Hence, some form of separation will be required. Major candidate approaches include various forms of micro liquid chromatography and capillary electrophoresis. Here also, rapid separation as well as minimalization of consumable materials are important target characteristics.

- Sample cleanup/preparation: The level of sample cleanup required prior to mass analysis will depend to a significant degree on the type of separation utilized. In any case, steps must be taken to minimize the concentration of common biological solution components that may have a deleterious impact on mass spectrometry. As noted earlier, salts are a major concern in this respect.

V. CONCLUSION

The successful development of a next-generation biological mass spectrometer coupled with on-line sample processing approaches is expected to provide an enhanced ABO identification capability with characteristics that are advantageous with respect to other identification approaches. Such characteristics include applicability to a broad range of ABOs with potentially a single assay format, capability to characterize unknown materials and perform a field-based update of its library, and the ready incorporation of a chemical detection/identification capability. To realize this potential, a number of key issues must be addressed as outlined to ensure that an optimum configuration mass spectrometer is developed in the near term based on the present state of the art technology and to ensure that avenues for further investigation and improvement are identified for future developmental efforts. This Workshop was conceived as a major element in the process to address these issues.

Blank

POLAR LIPID BIOMARKERS: A STRATEGY BEYOND FINGERPRINTING

Catherine Fenselau, Daniele Fabris, Yetrib Hathout, Yen Peng Ho, John Thomas
University of Maryland Baltimore County, Structural Biochemistry Center
Baltimore, MD 21250, USA
Robert Cotter
Johns Hopkins University, School of Medicine, Dept. of Pharmacology and Molecular Sciences
Baltimore, MD 21205, USA
Wayne Bryden
Johns Hopkins University, Applied Physics Laboratory
Laurel, MD 20713, USA

Earlier work has demonstrating that bacteria, viruses, fungi and mammalian tissue can be characterized by direct desorption of biomarkers into a mass spectrometer. This is being extended to the development of a smart field portable detection system.

The use of mass spectrometry to characterize microorganisms directly was first proposed in 1975 (1), in a report in which gentle heating was used to release large organic biomarkers. This landmark paper led to the development of two separate approaches, characterization by pyrolysis and characterization by interrogation of desorbed biomarkers. The latter has benefited dramatically from the development of techniques that ionize samples in the solid or liquid phase.

Plasma desorption, laser desorption and fast atom bombardment were all reported in the mid-1980's (2-4) to provide biomarkers from gram negative bacteria and from mammalian tissue. Fast atom bombardment was used to develop mass spectrometric strategies to characterize phospholipids and other polar lipids released from the outer membranes of gram positive and gram negative bacteria (5). Variations in phospholipid polar head groups and fatty acids, reflected in variations in molecular weights and fragment ion masses, provided profiles that could readily distinguish species manually or by computer-supported library searching (6).

The capabilities of tandem mass spectrometry have been shown to allow an interrogative or interpretive approach to be taken as well (7). All phospholipids undergo fragmentation that separates their polar head groups from the more variable fatty acids. Consequently, tandem mass spectrometry techniques allow polar head groups to be identified more reliably and to be quantified as part of identification. Constant neutral loss and other tandem scans provide improved sensitivity and allow biomarkers to be detected that occur in minor amounts or with overlapping masses (7). Some minor phospholipids occur rarely, and thus are excellent markers for particular species of bacteria. Quantitation from constant neutral loss spectra has been used to determine relative amounts of phosphatidylethanolamine and phosphatidylglycerol in *B. subtilis* and *E. coli* grown in different media and at different temperatures. Examples of these ratios obtained using fast atom bombardment on a JEOL four sector tandem instrument, are shown in Figure 1.

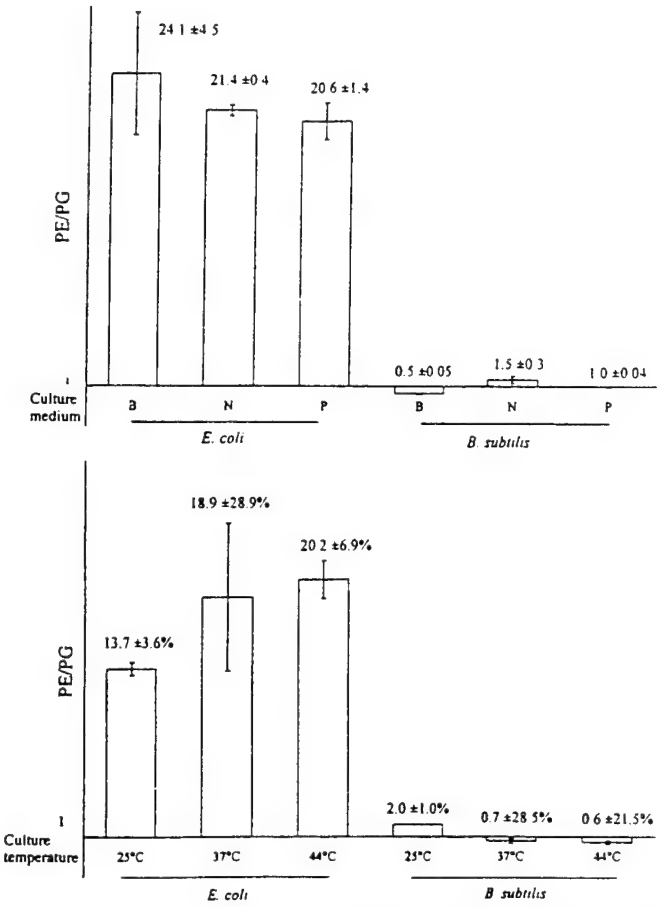


Figure 1. Ratios of phosphatidylethanolamine (PE) and phosphatidylglycerol (PG) in bacteria grown (top) in bovine brain heart infusion (B), nutrient broth (N) and vegetarian recipe (P) (bottom) at 25°C, 37°C and 44°C. Cells were harvested after 24 hours of growth.

When the full range of polar lipids can be detected, rapid broad classifications can be made based on the dominance of phosphatidylethanolamine in gram negative bacteria, the relatively high abundance of phosphatidylglycerol in gram positive bacteria, phosphatidylinositol in yeast, sulfonyl lipids in algae, and phosphatidylcholine in enveloped viruses (8). The plasma desorption spectrum of enveloped monkey cytomegalovirus is presented in Figure 2 (9). Phosphatidylcholine from the outer envelope can be seen clearly, along with sucrose from the preparation of the virus. The phosphatidylcholines detected reflect the distribution of mammalian host cells. Fast atom bombardment and multiple mass spectrometer techniques (ms/ms/ms) have been used to characterize the fatty acids in phosphatidylcholine species recovered from human immunodeficiency virus (10).

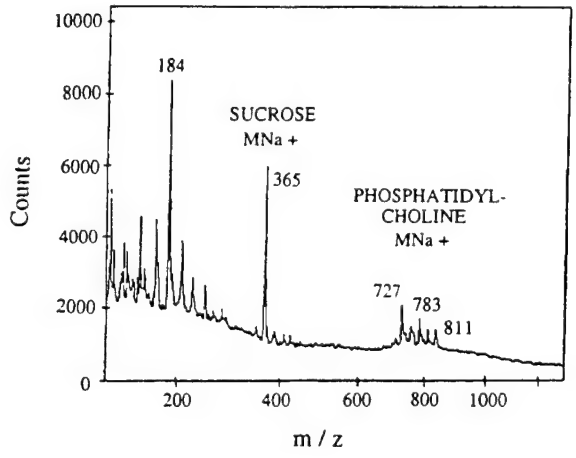


Figure 2. Plasma desorption spectrum of monkey cytomegalovirus.

The use of desorption techniques to analyze intact biomarkers to characterize microorganisms has been accepted worldwide, and fast atom bombardment in particular has been applied successfully in many labs in North America and Europe. Currently, electrospray ionization is being exercised for biomarker analysis (11) and matrix assisted laser desorption (MALDI) is extending the earlier work with laser desorption (12-16). An interrogative biomarker strategy is currently being implemented by this team, exploiting matrix assisted laser desorption and time-of-flight mass spectrometry in a small portable instrument (tiny TOF) (17). A schematic representation of the system is shown in Figure 3.

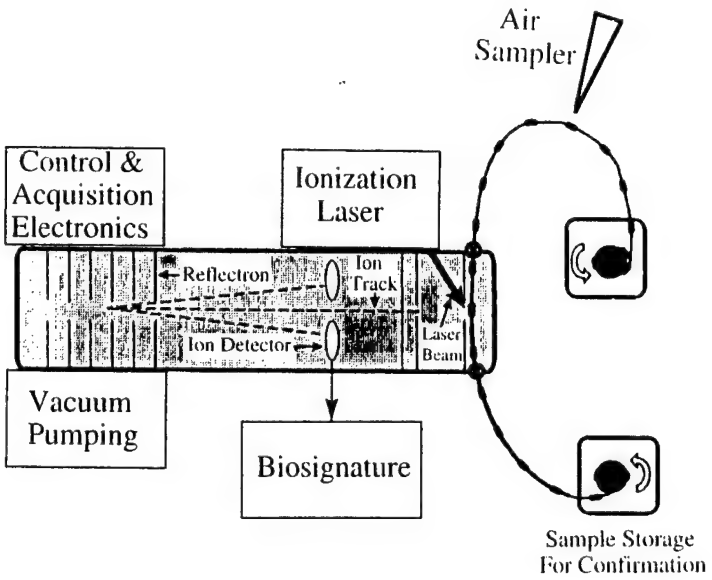


Figure 3. Instrument schematic for the tiny TOF-based field portable system (17).

The potential of this approach is illustrated in Figures 4 and 5. The mass spectra in Figure 4 shows phospholipid biomarkers desorbed by MALDI TOF and MALDI FTMS from the gram-negative bacteria *E. herbicola*, *B. thuringiensis* and *B. subtilis,* variant niger (*B. globigii*) spores have also been characterized by unique sets of polar lipids. Figure 5 presents the MALDI spectrum of an unfractionated aliquot of growth medium containing bacteriophage MS2 virus and the *E.coli* host cells used to grow the virus. Under conditions optimized for protein analysis, the coat protein of the virus provides the dominant signal in the spectrum, at m/z 13787 (calculated 13773 Da; $(M+2\ Na)^+$). Based on the concentration determined for plaque forming units, this signal represents approximately 1 fmol. of protein in the mixture. Thus proteins and polar lipids can be used in tandem as biomarkers to characterize unknown microorganisms, extending both the range of analytes and also the reliability of the analysis.

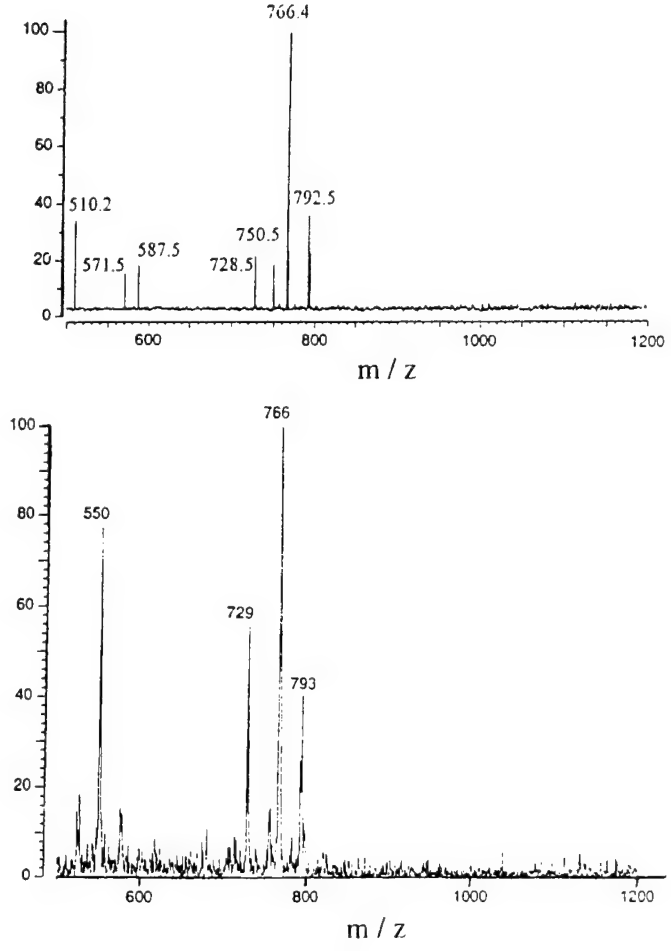


Figure 4. MALDI spectra of *E. herbicola* (top) measured on an FTMS and (bottom) measured on a TOF instrument.

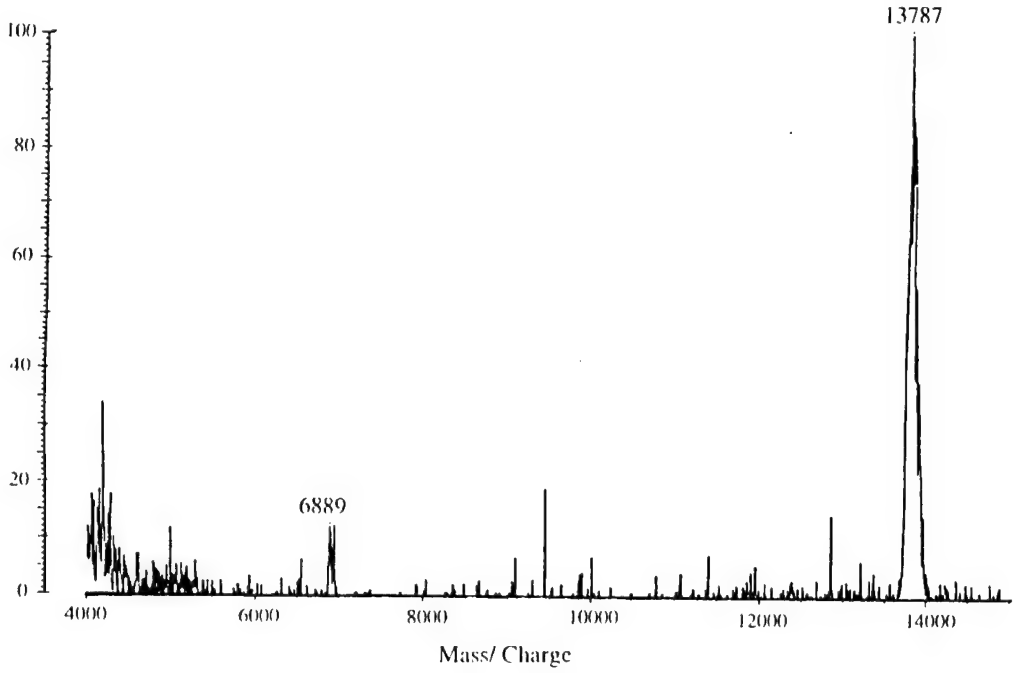


Figure 5. MALDI TOF spectrum of an aliquot from a culture of the bacteriophage MS2 virus in *E.coli*.

REFERENCES

1. Anhalt, J.P. and Fenselau C. Identification of Bacteria Using Mass Spectrometry. Anal. Chem. 47: 219-225 (1975).
2. Heller, D.N., Fenselau, C., Cotter, R.J., Demirev, P. Olthoff, J., Honovich, J., Uy, M., Tanaka, T., Kishimoto, Y. Mass Spectral Analysis of Membranes and Cell Wall: Comparative Selectivity in Desorption Techniques. The 34th Annual Conference on Mass Spectrometry and Allied Topics, June 8-13, 1986 Cincinnati, Ohio pp. 872.
3. Heller, D.N., Fenselau, C., Cotter, R.J., Demirev, P., Olthoff, J.K., Honovich, J., Uy, M. Tanaka, T., Kishimoto, Y. Mass Spectral Analysis of Complex Lipids Desorbed Directly from Lyophilized Membranes and Cells. Biochem. Biophys. Res. Commun. 142: 194-199 (1987).
4. Heller, D.N., Cotter, R.J., Fenselau, C., Kishimoto, Y., and Uy, O.M. Desorption of Ions from Rat Membranes: Selectivity of Different Ionization Techniques. Biomed. Environ. Mass Spectrom. 18: 1037-1045 (1989).
5. Heller, D.N., Cotter,f R.J., Fenselau, C., and Uy, O.M. Profiling of Bacteria by Fast Atom Bombardment Mass Spectrometry. Anal. Chem. 59: 2806-2809 (1987).
6. Platt, J., Uy, O.M., Heller, D.N., Cotter, R.J. and Fenselau, C. Computer Based Library Matching and Multivariant Analysis of Desorption Mass Spectra of Micro-organisms. Anal. Chem. 60: 1415-1419 (1988).
7. Heller, D.N., Murphy, C.M., Cotter, R.J., Fenselau, C. and Uy, D.M. Constant Neutral Loss Scanning for the Characterization of Bacterial Phospholipids Desorbed by Fast Atom Bombardment. Anal. Chem. 60: 2787-2791 (1988).
8. Fenselau, C. In: Mass Spectrometry for Characterization of Microorganisms: An Overview. (C. Fenselau, ed), ACS Symposium Series 541, American Chemical Society, Washington, D.C., Chapter 1, pp. 1-7 (1994).
9. Heller, D.N. personal communication 1997.
10. Bryant, D.K., Orlando, R.C., Fenselau, C., Sowder, R.C. and Henderson, L.E. Four Sector Tandem Mass Spectrometric Analysis of Complex Mixtures of Phosphatidycholines Present in Human Immunodeficiency Virus Preparation. Anal. Chem. 63:1110-1114 (1991).
11. Biological and Biotechnological Applications of ESI-MS (P. Snyder, ed.), ACS Symposium Series 619, American Chemical Society, Washington, DC pp. 424-431 (1996).
12. Parker, C.E., Papac, D.I. and K.B. Tomer. Monitoring Cleavage of Fusion Proteins by Matrix-Assisted Laser Desorption Ionization/Mass Spectrometry: Recombinant HIV-1_{111B}p26. Anal. Biochem. 239: 25-34 (1996).
13. Cain, T.C., Lubman, D.M. and Weber, W.J. Differentiation of Bacteria using protein Profiles from Matrix-Assisted Laser Desorption/Ionization Time-of-flight Mass Spectrometry. Rapid Communications in Mass Spectrometry 8: 1026-1030 (1994).
14. Claydon, M.A., Davey, S.N., Edwards-Jones, Valerie, Gordon, D.B. The rapid identification of intact microorganisms using mass spectrometry. Nature Biotechnology 14: 1584-1586 (1996).
15. Krishnamurthy, Thaiya and Ross, P.L. Rapid Identification of Bacteria by Direct Matrix-assisted Laser Desorption/Ionization Mass Spectrometric Analysis of Whole Cells.Rapid Communications in Mass Spectrometry 10: 1992-1995 (1996).
16. Holland, R.D., Wilkes, J.G., Rafii, F., Sutherland, J.B., Persons, C.C., Voorhees, K.J. and Lay, J.O., Jr. Rapid Identification of Intact Whole Bacteria Based on Spectral patterns using Matrix-assisted Laser Desorption/Ionization with Time-of-flight Mass Spectrometry. Rapid Communications in Mass Spectrometry 10: 1227-1232 (1996).
17. Bryden, W.A., Benson, R.C., Ecelberger, S.A., Phillips, T.E., Cotter, R.J. and Fenselau, C. The Tiny-TOF Mass Spectrometer for Chemical and Biological Sensing. Johns Hopkins APL Technical Digest 16: 296-309 (1995).

Blank

BACTERIAL IDENTIFICATION BY MASS SPECTROMETRY

*Thaiya Krishnamurthy**
US Army Edgewood RDE Center, Aberdeen Proving Ground, MD 21010-5423 USA

Jimmy Eng and John Yates
Molecular Biotechnology, University of Washington, Seattle, WA 98195-7730 USA

Michael Davis and Terry Lee
Beckman Research Institute, City of Hope, Duarte, CA 91010-0269 USA

Uma Rajamani, Shengjun Wang, and Karen Heroux
Geo-Centers, Gunpowder Branch, P.O. Box 68, Aberdeen Proving Ground, MD 21010

Terry Abshire, John Ezzell, and Erik Henchal
U.S. Army Medical Research Institute for Infectious diseases, Ft. Detrick, Frederick, MD

ABSTRACT

Unique, direct, rapid, sensitive and selective mass spectrometric methods have been developed for the identification of a bacterial pathogen by its strain. The procedure requires negligible sample preparations prior to mass spectral analysis and can be completed in less than 10 minutes. Intact bacterial cells were mixed with uv absorbing matrix and subjected to MALDI-TOF-MS analysis. Alternatively, the components of the bacterial suspensions were separated over a reverse phase fused silica capillary column and subjected to on-line Electrospray ionization and detected by Ion-trap mass spectrometry. In both instances, biomarkers specific to the individual organism were identified and applied for analyzing bacteria present in samples. Mixture analysis was also possible under either one of these experimental conditions. An algorithm for automated identification of the bacteria from the acquired mass spectrum was also developed and utilized for sample analysis. These procedures have enormous application potential in field operations as well as in monitoring agricultural products, meats, water resources, health institutions and products, and environmental and biological samples.

INTRODUCTION

Proteinaceous toxins, pathogenic bacterial cells and viral particles have been known to cause health hazards to humans and hence pose a great danger during field operations. The analytical techniques and procedures for the identification of the threat should be universal so that they can be applied not only for identification of known threat agents but also for characterization of unknown ones. Hence, we applied the state-of-the-art Electrospray ionization(ESI)- and matrix assisted laser desorption ionization(MALDI)- mass spectrometric techniques and small commercial bench-top mass spectrometers to develop analytical procedures for bacterial identification which has great application potential in laboratories as well as during field operations.

Biopolymers such as lipids, carbohydrates, proteins and DNA present in all bacterial cells have been known to be potential biomarkers for any individual organism.[1,2] We selected to investigate the bacterial protein biomarkers since the bacterial proteins provide indirect information on the genome of the investigated organisms and comprise as much as 60% by weight of the dry bacterial cells.[2] Hence we projected that the bacterial proteins could be utilized to distinguish the pathogenic bacteria from its corresponding non-pathogenic species.

In addition, the same techniques and possibly procedures with slight modifications can be applied for identifying other biological agents such as protein toxins and viral particles.

Individual bacterial pathogens (*Bacillus anthracis, Brucella melitensis, Francisella tularensis and Yersinia pestis*) could easily be distinguished from their corresponding MALDI-time-of-flight (TOF)-mass spectra of the cellular proteins.[3] The same procedure was applied for the clear distinction between individual members of *Bacillus cereus* group (*B.anthracis, B.thuringiensis* & *B.cereus*) and also from a non-pathogenic related species (*B.subtilis*).[3] Even though this approach fulfilled all requirements, the total analysis time (2-3 hr)[3] was unacceptable. This was rectified by subjecting the aqueous suspensions of intact whole bacterial cells directly to the MALDI-MS analysis and the entire process was completed with comparable results within 10 minutes.[4]

We also separated the bacterial suspensions in 0.1% aqueous trifluoroacetic acid (TFA) over a fused silica capillary reverse phase HPLC column. The separated components were directly introduced into the ESI source and the generated ions were detected by an Ion-trap mass spectrometer. The entire process could be completed within 10 minutes and generated specific biomarkers can be applied for the identification of bacteria distinctly in samples. This methodology provides great advantage than our earlier procedures since more specific biomarkers can be generated in the future from the product ion (MS/MS) spectra of the bacterial proteins.

There are other methods reported in the literature based on the analysis of bacterial lipids,[5,6] proteins,[7] DNA,[8] and intact cells[9,10] of certain organisms. However, our approach was not only systematic but also generated ions with excellent sensitivity and low background noise indicating the sample preparation process was adequate. In addition, we have also tested the developed methodology for potential applications involving identification of cellular pathogens.

EXPERIMENTAL

The entire investigation was carried out using PerSeptive Voyager-Elite or Voyager-DE MALDI-TOF-MS instruments (PerSeptive Biosystems, Framingham, MA) and Finnigan-MAT LCQ tandem mass spectrometer (San Jose, CA). Experimental details, including MALDI-MS analysis of the protein extracts and whole cells are described elsewhere in the literature.[3,4]

AUTOMATED BACTERIAL IDENTIFICATION: The individual acquired mass spectral file was converted into the corresponding ASCII file. Relevant information such as the genus, species and strain of the investigated bacteria, mass range and origin of the generated spectra, etc., are included. The full spectrum is read, reduced by selecting the maximum intensity peaks at every 10 amu range, and the reduced spectrum with annotations was stored. Sample input file, generated by a similar process is subjected to comparison with the library spectra. During the process, the sample spectrum loops through each library file and cross correlation against each library file is performed and the individual scores are stored and sorted. Ten library spectra with highest scores assigned during the matching are displayed with corresponding scores. Comparative display of the sample spectrum and the library spectrum with the highest score is also accomplished.

LC-ESI/MS ANALYSIS OF INTACT WHOLE CELL BACTERIA: Experimental procedures can be found elsewhere in detail.[3,4]

RESULTS & DISCUSSION

Intact bacterial cells suspended in 0.1% TFA were mixed with sinapinic acid and subjected to MALDI-TOF-MS analysis. The entire process could be completed within ten minutes and absolutely no sample processing was required prior to analysis. Ions generated under MALDI-MS conditions were accelerated, after a time lapse of 150-200 nsec (delayed

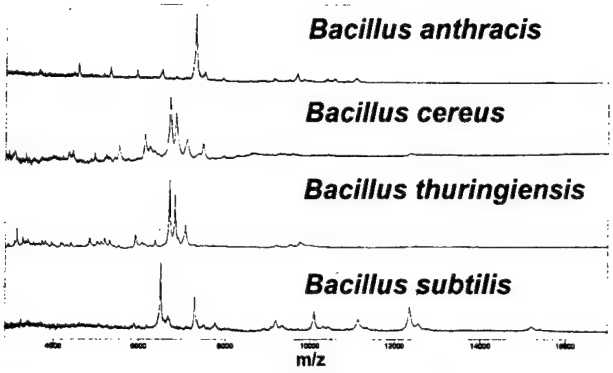


Figure 1. MALDI-MS analysis of *Bacillus* species.
Intact cell suspensions analyzed using a bench top MALDI-TOF-MS system. Linear-delayed extraction data. Sinapinic acid matrix

extraction), through the flight tube and detected under linear mode. Data generated in a bench top mass spectrometer (Voyager-DE), illustrated in this article, was quite comparable to the data from a research type instrument (Voyager-Elite). Excellent resolution of the peaks and hence more accurate mass assignments could be achieved, even without the internal calibrants. Distinction of the pathogens (*B.anthracis, B.melitensis, F.tularensis and Y.pestis*) could easily be accomplished from the corresponding MALDI-MS spectra generated during the analysis of the intact bacterial cells. Similarly, differentiating between the members of the *B.cereus* sub-species (*B.cereus, B.anthracis* & *B.thuringiensis*) and a non-pathogenic species (*B.subtilis*) can easily and directly be achieved (Figure 1). In most instances, crude samples can be analyzed directly and on few

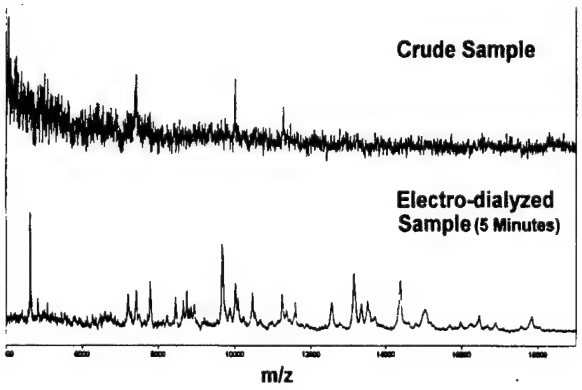


Figure 2. Electro-dialysis of a crude bacterial sample.
Sinapinic acid matrix.

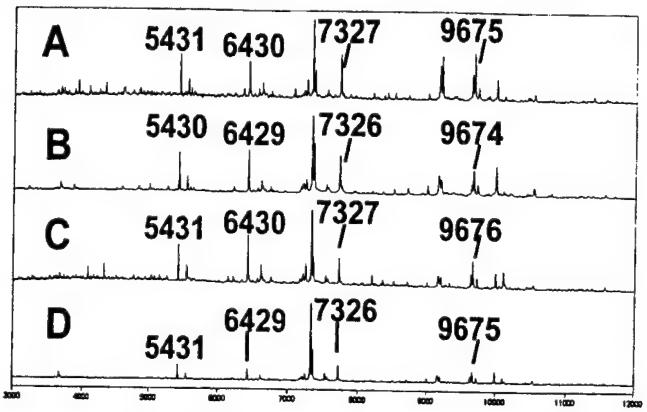


Figure 3. Effect of growth conditions on observed biomarkres of *Bacillus cereus,*12826.
A, Blood chocolate agar; B, Brucella agar; C, Blood TSA agar; D, TSA agar.

occasions desalting of the sample either by water wash or electro-dialysis of the sample was required. Both of these clean-up processes could be completed within 5 minutes (Figure 2).

Comparable biomarkers were generated for the same bacterial cells grown under different culture media (Figure 3). Similarly, regardless of the status of the cell, such as vegetative Vs sporulated or virulent Vs non-virulent, similar biomarkers were detected. Some of the biomarkers observed during the MALDI-MS analysis of the intact *Bacillus* cells are listed in Table 1. Artificial mixtures containing 1-4 bacterial components were also analyzed without any separation or processing prior to MS analysis. Individual bacteria could be identified from the recorded MALDI-MS spectrum (Figure 4). The designated biomarkers could be generated repeatedly from the corresponding bacterial cells when they were analyzed using 5 instruments from 3 manufacturers. This clearly demonstrates the enormous application potential of the developed methodology for the identification of gram positive as well as gram negative bacteria in samples, easily and reproducibly with excellent sensitivity.

In order to automate the process, an algorithm was developed both for the storage of acquired standard spectra in a specified MS library and comparison of the sample spectrum with the library spectra automatically. The details of this process is described in the

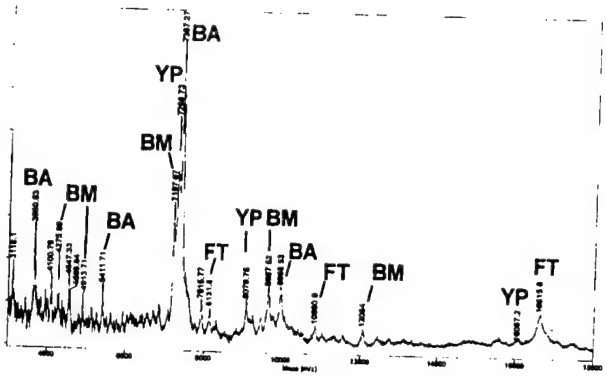


Figure 4. Bacterial mixture analysis by MALDI-MS technique.
Ba, *B.anthracis*; Bm, *B.melitensis*; FT, *F.tularensis*; YP, *Y.pestis*.

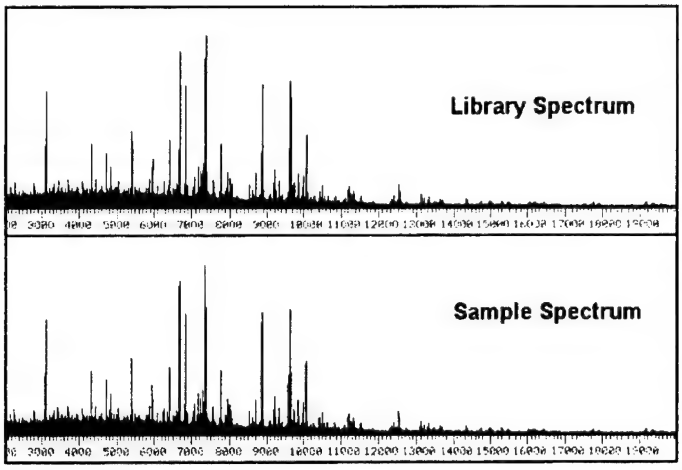


Figure 5. Identification of B.anthracis,Sterne in a sample by an automated library search routine

Table 1. Biomarkers for *Bacillus* Organisms

Organism	Genus	Species
B.anthracis	6845,7189,7389,7793,8877 9674,11573,12551	2715,2821,5560,5638,6632, 8880,10100
B.thuringiensis	6840,7187,7371,7755,8876 9647,11559,12525	3809,4438,8876,19070
B.cereus	6840,7168,7371,7773,7773 9648,11559,12514	4331,5798,6380
B.subtilis	7170,7780,11570	2786,2993,6167,7170,7780, 9146

Table 2. Biomarkers from ESI-MS Investigations

Organism	Genus	Species
B.anthracis	6429,7775,9188,9336, 9996	4231,4942,5835,5913, 6201,6683,7017,7558, 8465,8846,9524,10071, 11376
B.thuringiensis	7774,9160,9335,9994	2790,2805,3441,3798, 3850,5578,5616,6334, 6611,9670,9734
B.cereus	6429,7771,9187,9332, 9992	3937,5458,5549,6253, 6905,10507,11234
B.subtilis	7771,9187	6511,7301,7470,9062, 9891,11140,15201

experimental section. In most instances, the bacteria present in the samples were correctly identified at least by its genus and species and the score for the matched standard spectrum during the library search was higher than the other entries in the library. Comparative displays of the sample and library spectra origination during the searches are illustrated in Figures 5. All of the library spectra were generated in the linear delayed extraction mode using PerSeptive Voyager-Elite. In order to test the process more rigorously, the sample spectra in some cases were generated by analyzing the crude samples without any sample clean-up using a bench-top Voyager-DE mass spectrometer. Despite this major variation, identifications were correctly made for samples in most instances. Optimization of the search routines are presently underway to enable the identification of the bacteria up to its strain level.

Despite the success of the MALDI-MS methodology, we investigated an alternate method for identification of bacterial pathogens by a yet more specific tandem mass spectrometric methodology which can be completed within the allowable time frame. During this investigation, we injected (1 µl) of the bacterial suspension into a fused silica capillary reverse phase (C_8) HPLC column. The separated components were introduced on-line into the ESI source and ionized. The ions were focused and detected in an Ion-trap mass spectrometer. A mixture of specific peptides and proteins were detected when individual bacteria were analyzed. It was later deduced that the trifluoroacetic acid present in the solvent was sufficient to cause cell lysis and release the cellular proteins along with peptides. When the sample was reconstituted in 20% acetonitrile in 0.1% TFA, only proteins were detected

(Figure 6). Thus we could selectively utilize either protein or peptide biomarkers for our

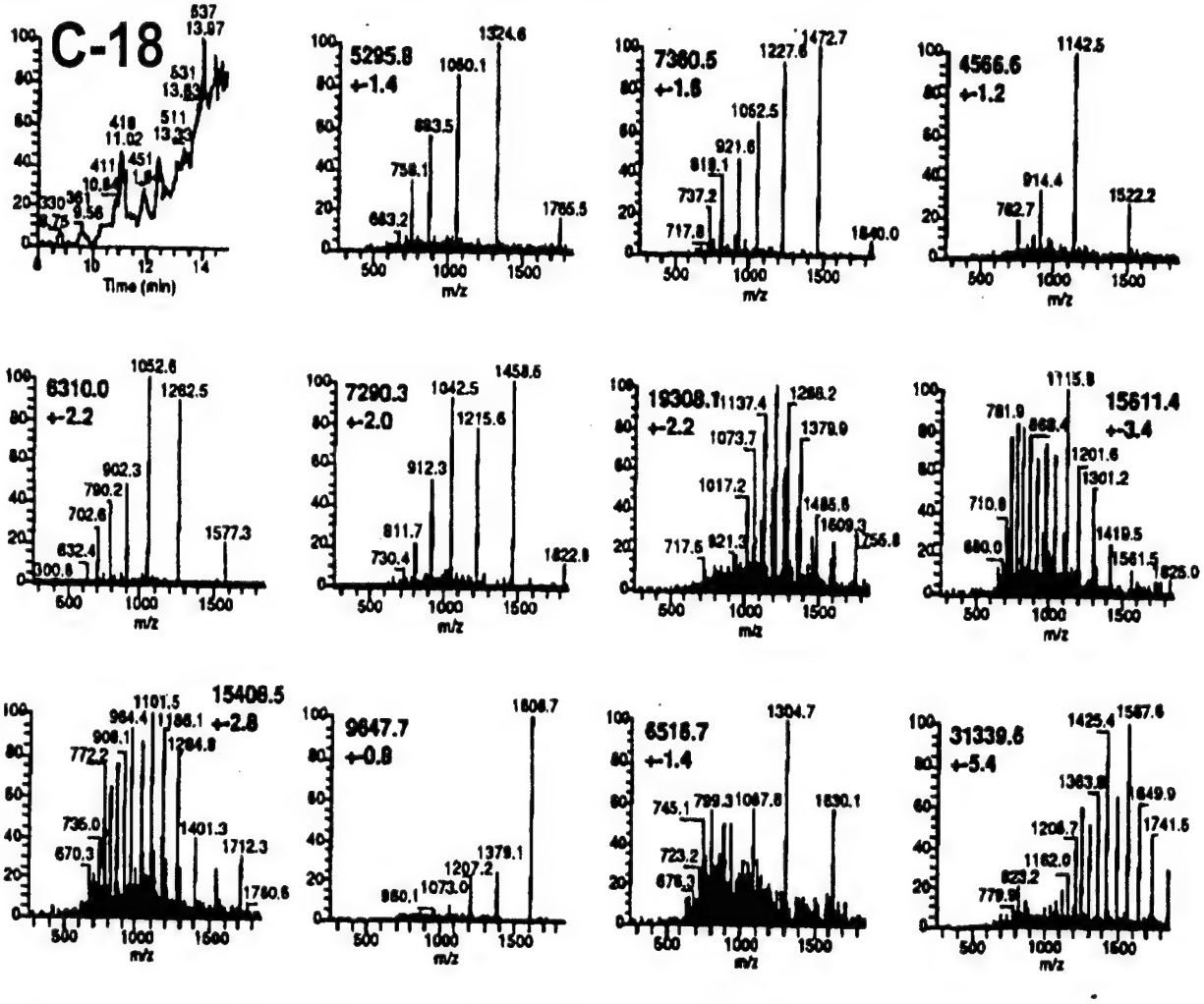


Figure 6. LC/ESI-MS analysis of intact *Yersinia pestis* cells.
Data was generated in an Ion-trap mass spectrometer. Protein separation was performed over a fused silica capillary reverse phase HPLC (C-18) column.

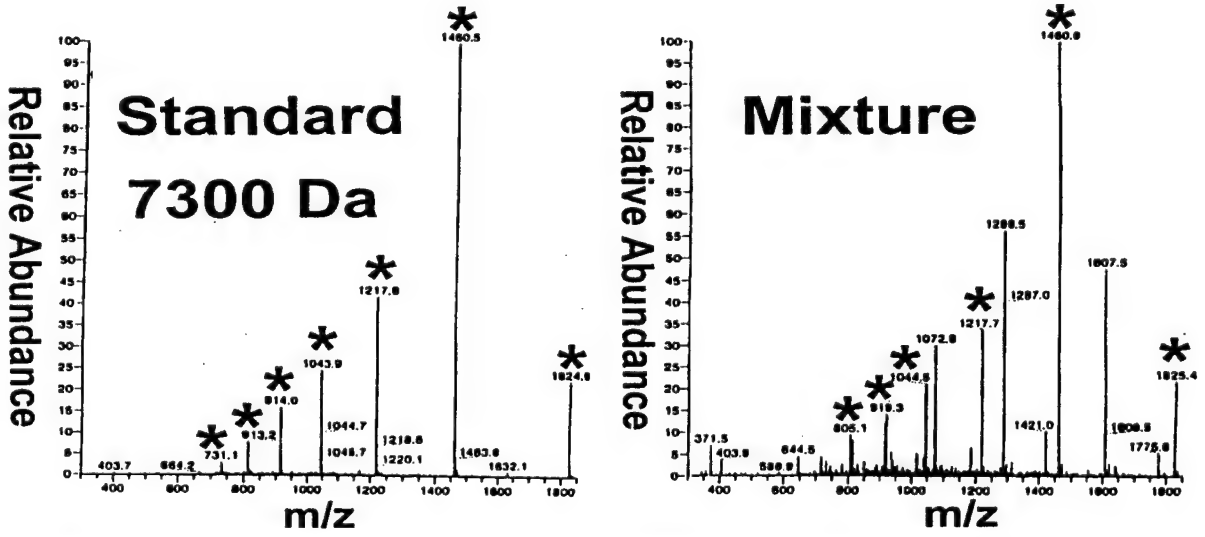


Figure 7. Comparison of ESI-MS spectra, standard Vs sample, of *B. Subtilis* During LC/ESI-MS analysis of a bacterial mixture.

investigations. A series of specific biomarkers for some pathogenic as well as non-pathogenic gram positive and negative bacteria were generated from the acquired LC/ESI-MS data (Table 2). All of these biomarkers could be generated reproducibly on repeated analyses, similar to the observations made during the MALDI-MS investigations. However, the observed protein biomarkers observed during the ESI- and MALDI- mass spectrometric analyses are different (Table 1 & 2). This is probably due to the ionization supression of less ionizable proteins during the analysis of the entire mixture containing numerous proteins, without further HPLC separations prior to the MS analysis. The differences observed between these two methods, despite the observed biomarkers are consistent and reproducible during any specific adopted method, are presently under investigation.

An artificial mixture containing equal amounts of *B.anthracis, B.thuringiensis, B.cereus,* and *B.subtilis* was made and analyzed by the same LC/ESI-MS procedure.The organisms were identified by comparing the acquired spectrum with the corresponding standard spectrum. Multiply charged ions of a biomarker for the organism was observed distinctly in the sample spectrum (Figure 7). Further investigations towards generating more specific biomarkers from the product ion spectra of these peptide and protein biomarkers are underway.

CONCLUSIONS

Protein biomarkers required for the automated identification of specific bacteria can be generated both by MALDI-MS and ESI-MS techniques in a reproducible manner. The developed algorithm could be successfully applied for the transfer of the acquired standard spectrum into a MS library and automated comparison of the sample spectrum with the library entries in order to identify the bacteria present in unknown samples.

REFERENCES

1. C. Fenselau, (Ed., 1994), Mass Spectrometry for the Characterization of Microorganisms, ACS Symposium Series, Volume 541.
2. M. Goodfellow and A.G.O'Donnell (Ed., 1993), Handbook of New Bacterial Systematics, Academic press, SanDiego, CA.
3. T. Krishnamurthy, P.L. Ross and U.Rajamani, Rapid Commun. Mass Spectrom., **10,** 883-888 (1996).
4. T. Krishnamurthy and P.L. Ross, Rapid Commun. Mass Spectrom., **10**, 1992-1996 (1996)
5. D.N. Heller, R.J. Cotter, C. Fensalau and O.M. Uy, Anal.Chem., **59**, 2806-2809 (1987)
6. P.B.W. Smith, A.P. Snyder and C.S. Harden, Anal.Chem., **67**, 1824-1830 (1995)
7. D..C. Muddiman et al., Anal. Chem., **68**, 3705-3712 (1996)
8. X. Liang, K. Zheng, M.G. Qian and D.M. Lubman, Rapid Commun. Mass Spectrom., **10**, 1219-1226 (1996)
9. R.D. Holland, J.G. Wilkes, F. Rafii, J.B. Sutherland, C.C. Pearsons, K.J. Voorhees and J.O. Lay, Jr., Rapid Commun. Mass Spectrom., **10**, 1227-1232 (1996)
10. M.A. Clayden, S.N. Davey, V. Edwards-Jones and D.B. Gordon, Nature Biotechnology, **14**, 1584-1586 (1996)

SEQUENCING PEPTIDES at the SUB-FEMTOMOLE LEVEL from COMPLEX MIXTURES

Jeffrey Shabanowitz and Donald F. Hunt
University of Virginia
Department of Chemistry
Charlottesville, Virginia, 22901, USA

ABSTRACT

Structural characterization of peptidic tumor antigens and identification of their protein source is an important step in designing new approaches to immunotherapy against cancer. Complex mixtures of peptides (>10,000) are presented by the MHC class I molecules to the immune system for screening to determine the efficacy of the cell. Mass spectrometry has successfully been combined with an immunological (cytotoxic T-cell) assay to identify more than two dozen tumor antigens to date. Some of these antigens appear to be post-translationally modified, illustrating an advantage of direct sequencing by mass spectrometry versus other strategies. In the absence of a biological assay, a subtractive approach must be employed to determine the intracellular protein differences between a diseased and healthy cell. This is the case for peptides presented to the immune system in association with the MHC class II pathway. These peptides play an important role in a number of autoimmune diseases such as arthritis, diabetes, and multiple sclerosis. Class II peptides are typically 15-30 residues long, contain multiple basic residues, and are derived from complex mixtures at low levels further challenging mass spectrometry. The ability to perform MS^n analysis has proven to be advantageous. These techniques are also applicable to the characterization of differences in the protein content of whole cells (normal vs. diseased) identified via 2D-Gel electrophoresis. The methodology developed to sequence peptides at the sub-femtomole level combines microcapillary HPLC or CE separation on-line with micro- or nano-ESI, and tandem mass spectrometry.

1. INTRODUCTION

To understand the importance of identifying peptide antigens it is necessary to describe how cells communicate their health status to the immune system. Most cells in the human body synthesize multiple copies of several thousand proteins each day. Additional proteins are manufactured if the cell is infected with a virus or has undergone transformation to a cancer cell. To tell the immune system what is being made at any particular time, a few copies of most proteins are processed into 8-12 amino acid fragments or peptides and are presented on the cell surface.

The methodology developed analyzes the mixture of 10,000 peptides presented to the immune system on the surface of cells and identifies the single molecule, antigen, that triggers the immune system to attack and kill viral, and bacterial agents including cancer. One copy of a foreign antigen is thought to be sufficient to stimulate an effective immune response. From $3x10^9$ cells (1 liter of cells), the quantity of peptide present at 1/copy per cell would be approximately 6 fmol. Identification of a single peptide antigen present at this level in a mixture with 10,000 other peptides represents a significant analytical challenge. The technology involves interfacing microcapillary chromatography and tandem mass spectrometry with a sensitive immunological assay to identify the target antigen associated with a particular disease state. Mass spectrometry is also employed to deduce the structure of the target antigen, modifications of which can then be used as a vaccine or therapeutic against the disease in question. More than two dozen antigens have been identified to date (1-9).

In the absence of a biological assay, a subtractive approach must be employed to determine the intracellular protein differences between a diseased and healthy cell. This is usually the case for peptides presented to the immune system in association with the MHC class II pathway. These peptides play an important role in a number of autoimmune diseases such as arthritis, diabetes, and multiple sclerosis. Class II peptides are typically 15-30 residues long, contain multiple basic residues, and are derived from complex mixtures at low levels further challenging mass spectrometry. The ability to perform multiple stages of mass spectrometry (MS^n) analysis will prove to be advantageous. These techniques are also applicable to the characterization of differences in the protein content of whole cells (normal vs. diseased) identified via 2D-Gel electrophoresis.

Recent advances in the field of mass spectrometry (MS) have made it even more powerful for the sequence analysis of peptides. First, the electrospray ionization source has been miniaturized. This improves the ion transmission efficiency into the MS thus lowering the limit of detection. Second, research efforts over the last decade have transformed the ion trap mass spectrometer into a versatile high-performance instrument for biological applications. Reported here is the combination of microcapillary chromatography (HPLC and CE) with a miniaturized sheathless electrospray ionization source to a quadrupole ion trap mass spectrometer (LCQ).

The ion trap MS has two main advantages over the triple quadrupole MS for sequencing biologically relevant peptides. First, the ion trap MS is capable of multiple stages of analysis (MS^n). Peptides containing basic residues in the middle of their sequence often do not fragment randomly resulting in ambiguities in the interpreted sequence. The LCQ has the ability to isolate a product ion after the first stage of dissociation and fragmenting it further. This process can be repeated several times, as long as there is enough ion current, until the complete sequence is deduced. The second advantage of ion traps is in detection limits. The ion trap MS stores all of the ions in the desired mass range and then scans them out; a quadrupole MS, on the other hand, is continuously scanning over the mass range so one ion at a time is allowed to pass and therefore detected. This difference in operation translates into greater duty cycles and lower detection limits for the ion trap.

The combination of the immunological assay and all of the technological advances (ESI and MS) have resulted in a powerful technique for sequencing biologically relevant peptides at the sub-femtomole level.

2. RESULTS AND DISCUSSION

The methodology has been applied to the identification of a murine alloreactive MHC class I T-cell epitope. This peptide structure could not be solved on a triple quadrupole mass spectrometer at the 10 fmol level. By using the LCQ, MS^n, and the miniaturized ESI source, an unambiguous amino acid sequence was determined. These results were achieved at the 100 attomole level, an improvement in sensitivity of 20x over previously reported peptide epitopes by tandem mass spectrometry (8).

To identify the mass of the alloantigen, an aliquot of an active fraction was analyzed with an on-line microcapillary column effluent splitter as in described previous reports. This device combines the mass spectrometric with the immunological analysis by directing 4/5 of the effluent into the mass spectrometer while the other 1/5 is simultaneously directed into 96 well microtiter plates to be used for the biological assay. Since both types of data are acquired as a function of time, the epitope reconstitution activity in the active wells of the microtiter assay can be compared to a peptide's ion abundance in the corresponding mass spectra. A candidate mass at *m/z* 444, which corresponds to a protonated molecular mass of 887, matched the biological profile best.

To determine the sequence of the candidate peptide, collision-activated dissociation (CAD) mass spectra generated from the $(M+2H)^{+2}$ ions at *m/z* 444 were acquired on a triple quadrupole mass spectrometer. In this tandem MS experiment, the doubly protonated peptide molecules at *m/z* 444 are mass selected in the first quadrupole. Peptide molecules of other *m/z* values are unstable in the electric fields and are pumped away by the vacuum. These ions are then transmitted through an octapole (the collision chamber) which has ca. 3 mTorr argon as the target gas. The protonated peptide molecules undergo multiple low energy collisions that cause the peptide molecules to fragment randomly at the various amide bonds in the peptide. This generates a series of fragment ions that differ in mass by a single amino acid residue. These peptide fragments are then mass analyzed with a quadrupole as they exit the collision cell. The observed fragmentation of the candidate peptide was generated from 10 fmols of sample as shown in Figure 1 (note: an almost identical spectra can be obtained using only ca. 200 attomoles on the LCQ). Interpretation of the data resulted in only for the partial sequence, VAF, to be

deduced. Assignment of the total peptide sequence proved difficult due to the absence of high mass fragment ions that contained the N-terminus (b type ions) or low mass fragment ions that contained the C-terminus (y type ions). In this case, the MS/MS data did not yield a complete an amino acid sequence and the presence of an unusual amino acid or a post-translational modification could not be ruled out.

Multiple stages (MS^n) of collisionally activated dissociation (CAD) can often help generate this additional sequence information. In this experiment, using an ion trap MS, a product ion from the first stage of fragmentation is isolated, again subjected to CAD, and the second generation product ions are interpreted. The LCQ was used to fragment the $(M+2H)^{+2}$ ions at *m/z* 444. The singly charged product ions at *m/z* 530 were retained in the ion trap while the other product ions were removed using resonance ejection. The isolated ions were induced to fragment further by resonance excitation. Mass spectra shown in Figure 2 were generated from ca. 100 attomoles of peptide sample. The observed fragmentation was sufficient to specify the C-terminal residues 6-8 of peptide 887 as HNX, where X is either leucine or isoleucine, which are isobaric and cannot yet be differentiated under low energy CAD conditions. Hence, residues 3-8 of peptide 887 are VAFHNX and through other MS data, residues 1-2 are deduced to be SV.

To establish the identity of the C-terminal amino acid as leucine or isoleucine, co-elution studies were performed. The biologically active fraction containing the candidate peptide at*m/z* 444 was analyzed by μHPLC-MS before and after being spiked with the synthetic peptide SVVAFHNL or SVVAFHNI. A single peak at *m/z* 444 was observed for the synthetic peptide SVVAFHNL which suggests it to be the correct candidate. The synthetic peptide, SVVAFHNL, was then tested at various concentrations and confirmed to sensitize target T-cells to lysis by the CTL clone 13 (8).

Ion trap mass spectrometry in combination with micro-electrospray and HPLC was used in the identification of a murine alloantigen. This method required 20x less material than the more common triple quadrupole mass spectrometer. This reduction in sequencing detection limits will allow a reduction in the number of cells required for immunological assays. By using less cells for this type of assay this method of finding peptide antigens will be more time and cost effective. The ion trap was able to generate MS^3 spectra giving three residues of the peptide antigen not found in the triple quadrupole data. The MS^n capabilities of the ion trap are even more important for the identification of class II antigens (class II peptides typically contain more basic residues making sequencing them more difficult). The ion trap MS instrument will facilitate the identification of class I and class II antigens.

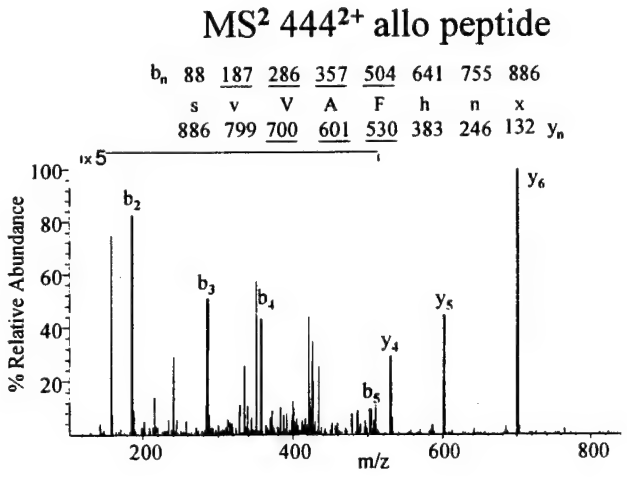


Figure 1

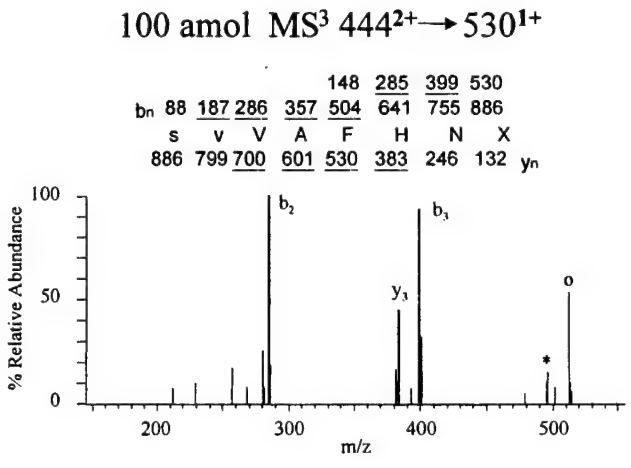


Figure 2

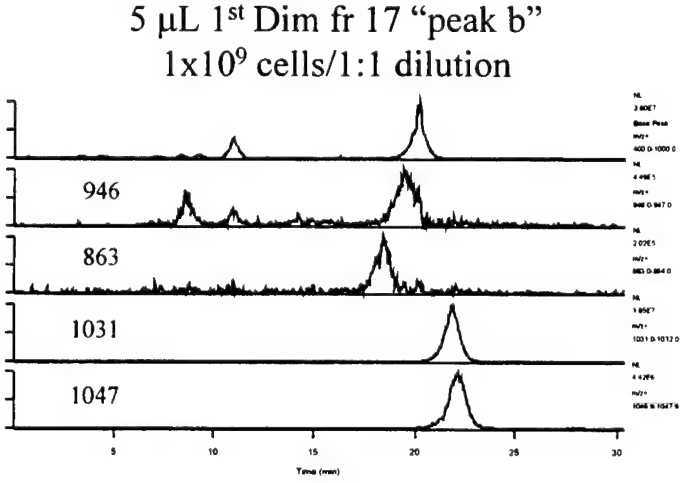
Another area of difficulty lies in the complexity of the peptide mixtures to be analyzed. Often, after two dimensions of HPLC separation, a one minute fraction may still contain hundreds of peptides which are not readily resolved using HPLC. Equally problematic is the possible presence of too many candidate peptides to sequence from the limited amount of material available. Furthermore, optimization of the ESI-MS performance requires the use of acids (i.e. acetic acid) which have lower chromatographic resolving capabilities. A different type of chromatography orthogonal to HPLC, such as capillary electrophoresis (CE), can be used to further differentiate the remaining candidate peptides (9).

CE offers very high chromatographic resolution but has other limitations for this application. First, the concentration limit of detection is dictated by the amount of sample that can be loaded, i.e. 10% of the column volume. Second, the capability and efficiency of interfacing of the CE to the ESI source and MS will determine the absolute detection limit. Suggested methods to overcome these shortcomings include focusing techniques (tITP) in association with preconcentration methods and microspray ionization. These techniques are developed, combined, and applied to the separation and detection of peptides derived from tumor cell extracts.

One method for reducing the concentration limit of detection is the use of a solid phase cartridge to preconcentrate the sample before the CE column. A polystyrene divinyl benzene (PSDVB) membrane was chosen because of less problems with reproducibility between cartridges, column backpressure, and ease of preparation. Miniaturization of electrospray ionization sources has shown a possible improvement in detection limit by 1-3 orders of magnitude. Of particular importance is the advent of the micro-electrospray ionization source (μ-ESI). The improvement results from sampling more of the μ-ESI plume by the MS. The plume is smaller due to the absence of sheath liquids and the lower flow rates exiting the spray emitter allowing a greater percentage of the analyte to enter the MS. Two different methods have been used to couple CE to ESI/MS. The first uses a liquid junction (usually a gap in the CE column contained in a metal tube). The second uses metalized needles where the CE termination voltage and μ-ESI voltage are applied as the liquid exits the CE capillary. Both approaches are problematic because of the possible presence of metal adduct ions obscuring the analyte of interest and the limited lifetimes of metal coated capillaries. A porous glass joint was chosen for the junction between the CE termination voltage and the application of μ-ESI voltage.

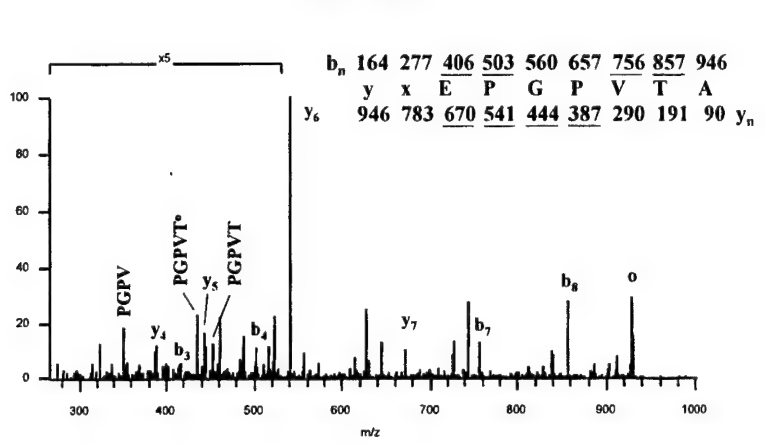
This report will demonstrate that a) CE can give complimentary information to that of our current HPLC methodology, b) we can use membrane preconcentration (mPC) techniques to overcome concentration detection limitations, and c) the absolute detection limits achieved are compatible with our samples. To test the complementarity of the separation techniques, a 1st dimension fraction from the extraction of human melanoma tumor cells, DM6, was used. From past work it has been observed that even after two dimensions of HPLC chromatography (1st dimension HFBA, 2nd dimension TFA) there are at least four candidate peptides to consider (two of which co-elute). Figure 3 depicts the ion chromatograms for each of these four candidate peptides obtained with CE. The active peptide found previously using μ-HPLC, at m/z 946, has now been separated from each of these other candidate peptides. The peptides that co-eluted using HPLC were 946 (the active one) and 1031.These peptides were baseline resolved without any effort optimizing the CE separation. Using CE as a 2 nd dimension separation technique should help in the narrowing of possible candidate peptides for sequence analysis.

Figure 3



CE is inherently limited in concentration detection limits. To overcome this limitation we used a mPC cartridge in association with tITP. After loading the sample, the column is rinsed with ca. 10 column volumes of run buffer prior to loading an elution buffer. This demonstrates the ability to concentrate the sample on a membrane and subsequently elute it off. In Figure 3, 5 μL (3x the column volume) of a 1:1 dilution of an unconcentrated HPLC fraction from 10^9 tumor cells was loaded. This represents ca. 250 amol of peptide 946 with a S/N ration of ca. 10. In addition, MS/MS spectra for all of the candidate masses were obtained in one run from only 1 uL of sample (shown in Figure 4 is the MS/MS of the active peptide 946).

Figure 4



The use of the porous glass joint has improved the limit of detection by approximately five times. Porous glass joints represent an improvement over the other liquid junctions tried by providing a more stable reproducible electrospray with a lower background. Capillary zone electrophoresis provided the necessary orthogonal separation to HPLC. As a result, detection limits are improved to less than 50 amol of peptide in both MS and MS^n modes of operation. This should allow the total number of cells needed to be reduced by a factor of 10 and make the analyses more routine and cost/time effective.

3. CONCLUSIONS

A technique to identify individual peptide epitopes found in mixtures estimated to contain at least 10,000 species has been described (1,3). Simultaneous immunological and mass spectrometric analysis was the most direct means to accomplish this objective. The analysis of MHC-Class I associated peptides using tandem mass spectrometry allowed for the identification of antigenic peptides without prior knowledge of the source proteins. This has important implications because prediction of potential antigens using the other approaches has been difficult and the DNA may not correctly identify the naturally occurring peptide sequence. Some of the antigens appear to be post-translationally modified illustrating an advantage of direct sequencing by mass spectrometry versus other strategies (6,7). This approach has proven to be useful for the identification and complete sequencing of individual peptides associated with melanoma and colon cancers, with malaria, and with transplantation rejection (1-2,4-6). These peptides are being used as drug candidates and provide leads in the discovery of new disease-associated antigens. The methodology has also been applied to the determination of antigenic peptides recognized by both MHC class I and class II-restricted T cells derived from virus or bacteria-infected cells and cells involved in autoimmune disease and in bone marrow and organ transplant rejection. This technology has been further enhanced by the application of microcapillary HPLC and/or CE separation techniques on-line with low-flow nanospray ESI and ion trap tandem mass spectrometry to sequence these biologically active peptides at the attomole level (8,9)

4. REFERENCES

1. AL Cox, J Skipper, Y Chen, RA Henderson, TL Darrow, J Shabanowitz, VH Engelhard, DF Hunt, CL Slingluff. Identification of a Peptide Recognized by Five Melanoma-Specific Human Cytotoxic T Cell Lines. Science 264: 716-719, 1994.

2. JMM den Haan, NE Sherman, E Blokland, E Huczko, F Koning, JW Drijfhout, J Skipper, J Shabanowitz, DF Hunt, VH Engelhard, E Goulmy. Identification of a Graft Versus Host Disease-Associated Human Minor Histocompatibility Antigen. Science 268: 1476-1480, 1995.

3. VH Engelhard. Structure of Peptides Associated with Class I and Class II MHC Molecules. Annu Rev Immunol 12: 181-207, 1994.

4. W Wang, LR Meadows, JMM den Haan, NE Sherman, Y Chen, E Blokland, J Shabanowitz, AI Agulnik, RC Hendrickson, CE Bishop, DF Hunt, E Goulmy, VH Engelhard. Human H-Y: A Male-Specific Histocompatibility Antigen Derived from the SMCY Protein. Science 269: 1588-1590, 1995.

5. AYC Huang, PH Gulden, AS Woods, MC Thomas, CD Tong, W Wang, VH Engelhard, G Pasternanck, R Cotter, D Hunt, et al. The Immunodominant Major Histocompatibility Complex Class I-Restricted Antigen of a Murine Colon Tumor Derives from an Endogenous Retroviral Gene Product. Proc Natl Acad Sci USA 93: 9730-9735, 1996.

6. L Meadows, W Wang, JMM den Haan, E Blokland, C Reinhardus, JW Drijfhout, J Shabanowitz, R Pierce, AI Agulnik, CE Bishop, et al. The HLA-A*0201-Restricted H-Y Antigen Contains a Posttranslationally Modified Cysteine that Significantly Affects T Cell Recognition. Immunity 6: 273-281, 1997.

7. JCA Skipper, RC Hendrickson, PH Gulden, V Brichard, A Van Pel, J Shabanowitz, T Wolfel, CL Slingluff, T Boon, DF Hunt, VH Engelhard. An HLA-A2 Restricted Tyrosinase Antigen on Melanoma Cells Results from Post-Translational Modification and Suggests a Novel Processing Pathway for Membrane Proteins. J Exp Med 183: 527-534, 1996.

8. RC Hendrickson, RE Settlage, J Shabanowitz, J Briggs, LA Sherman, DF Hunt. Identification of an Alloreactive Peptide Recognized by a Murine CD-8^+ T Cell by Ion Trap Mass Spectrometry at the Attomole Level. In preparation.

9. RE Settlage, PS Russo, J Shabanowitz, DF Hunt. Characterization of a u-ESI Source for Coupling Capillary Electrophoresis and Mass Spectrometry: Sequence Determination of Tumor Peptides at the Attomole Level. J. Chromatography, accepted.

CARBOHYDRATES AND MICROBIAL SURFACES

Bruce Reinhold, Walter Muehlecker, and Vernon Reinhold
Boston University School of Medicine
Mass Spectrometry Resource
Boston, Massachusetts

ABSTRACT

The gap between the ability of microorganisms to mutate around drug-resistant strains and man's ability to counter them is widening. Nearly half of the infections patients obtain in hospitals are caused by four bacteria: staphylococci, enterococci, pneumococci, and pseudomonades, and various strains of these bacteria have become resistant to all but a few antibiotics. By simple selective processes, bacteria, viruses and fungi have evolved ways to survive in the presence of drugs. Ironically, the success and widespread use of antibiotics contribute to drug resistance for long drug exposures select for evading properties, giving microorganisms an unparalleled tenacity for life. Our society faces a potentially endless race to outsmart evolving microbes, which take a while to develop new resistance's, but they certainly will. Recent findings have shown that many initial host cell contacts involve carbohydrate residues and our lack of a detailed understating of these structures has created a biological "ozone hole" which could offer unique opportunities for control and modulating the immune system.

Lipopolysaccharides (LPS) are a family of glycolipids present in the outer membrane of Gram-negative bacteria and are essential for the physical integrity of the membrane. These surface antigens exhibit host cell molecular mimicry, contribute to bacterial virulence, pathogenicity, and posses potent immunomodulating characteristics. Recent reports suggest glycolipid antigens can be presented to T cells by human CD1 molecules and their isotypes (1-3). Mycobacteria lysates and T cell clones derived from the blood of human subjects suffering from leprosy or tuberculosis have been fractionated to provide lysates that stimulate different T cell lines. Detailed structures features have been ascertained by electrospray using triple quadrupole and ion trap instruments. Mycolic acids and glycolipids from mycobacterial extracts were identification and structural relationships ere investigated to probe the broader question of antigen specificity in the CD1 restricted response.

(1) Porcelli et al., nature 360, 593 (1992).
(2) Sieling et al., Science 269, 227 (1995).
(3) Beckman et al., Nature 372, 691 (1994).

Blank

FLEXIBLE, HIGH-THROUGHPUT PATHOGEN DETECTION BY TIME-OF-FLIGHT MASS SPECTROMETRY

Christopher H. Becker
GeneTrace Systems Inc.
333 Ravenswood Avenue
Menlo Park, CA 94025

ABSTRACT

Parallel paths are being pursued for the analysis of DNA by matrix-assisted laser-desorption/ ionization (MALDI) time-of-flight mass spectrometry with applications to pathogen detection. We are refining analytical procedures for DNA analysis using 3-hydroxypicolinic acid as the matrix. Both sequencing and sizing assays are analyzed by these methods. For Sanger sequencing protocols starting with realistic amounts and types of target DNA (e.g., subpicomolar quantities of plasmid DNA), the state-of-the-art of our capabilities is reflected in the analysis of sequencing products extending to approximately 100 nucleotides. For sizing applications, a variety of genetic defect and genomic marker assays have been examined including those for microsatellites and single nucleotide polymorphism. At the same time, we are developing novel matrix approaches which will extend the accessible mass range for DNA.

1. INTRODUCTION

Until recently even very short pieces of DNA consisting of only a few nucleotides were beyond the capabilities of mass spectrometry for detection and identification. In 1988, a new mass spectrometric method called matrix-assisted laser-desorption/ionization (MALDI) was introduced for proteins.[1] This technique immediately expanded the mass range and sensitivity for detection of proteins. However, success using this new technique for fragile DNA molecules was still elusive. As recently as 1992, the largest DNA molecule that had been successfully analyzed using MALDI combined with time-of-flight (TOF) mass spectrometry, was a thymidine 18-mer.[2,3] Virtually no DNA of mixed sequence could be detected. When a suitable matrix compound (3-hydroxy-picolinic acid) for DNA mass spectrometry was discovered in our laboratory[4], a new capability opened for DNA analysis. DNA molecules could now be gently ionized and transferred into the gas phase without significant fragmentation. Further improvements quickly followed. Mixed-based oligonucleotides now are separated and detected intact with sensitivities in the low femtomole range.

2. RESULTS

2.1 APPLICATIONS FOR DNA ANALYSIS

At the current state-of-the-art, a wide variety of applications are becoming apparent including those relevant to potential biological warfare pathogens. Sequencing by the enzymatic (Sanger) approach using dideoxynucleotides for chain termination is the most general and powerful strategy.[5] Using this approach on practical amounts and types of target DNA, such as templates at subpicomole levels, high fidelity sequences are routinely obtained in our laboratory, such as shown in Fig. 1. Sequencing applications include pathogen detection, gene discovery, gene expression level analysis, exon sequencing, mutation screening, and *de novo* genomic sequencing. Current read lengths provide sufficient information for screening cDNA libraries. Mass spectrometrically derived expressed sequence tags (ESTs) can be used to

screen existing databases and identify novel gene products. Thousands of cDNAs from a given library can be sampled in an automated, high-throughput fashion to provide gene expression levels. Initial sequencing results using software developed in our lab have yielded base calling accuracy >98% with this number expected to increase as the recently developed algorithms are refined. The normally small size of exons make them excellent candidates for sequencing by mass spectrometry both for *de novo* identification and for mutation studies. One or more exons extracted from thousands of samples can be quickly screened to create a mutation profile for a given gene. Our laboratory is also currently examining the possibility of *de novo* genomic sequencing where mass spectrometric analysis is coupled with the use of transposon-based methods for cloning.

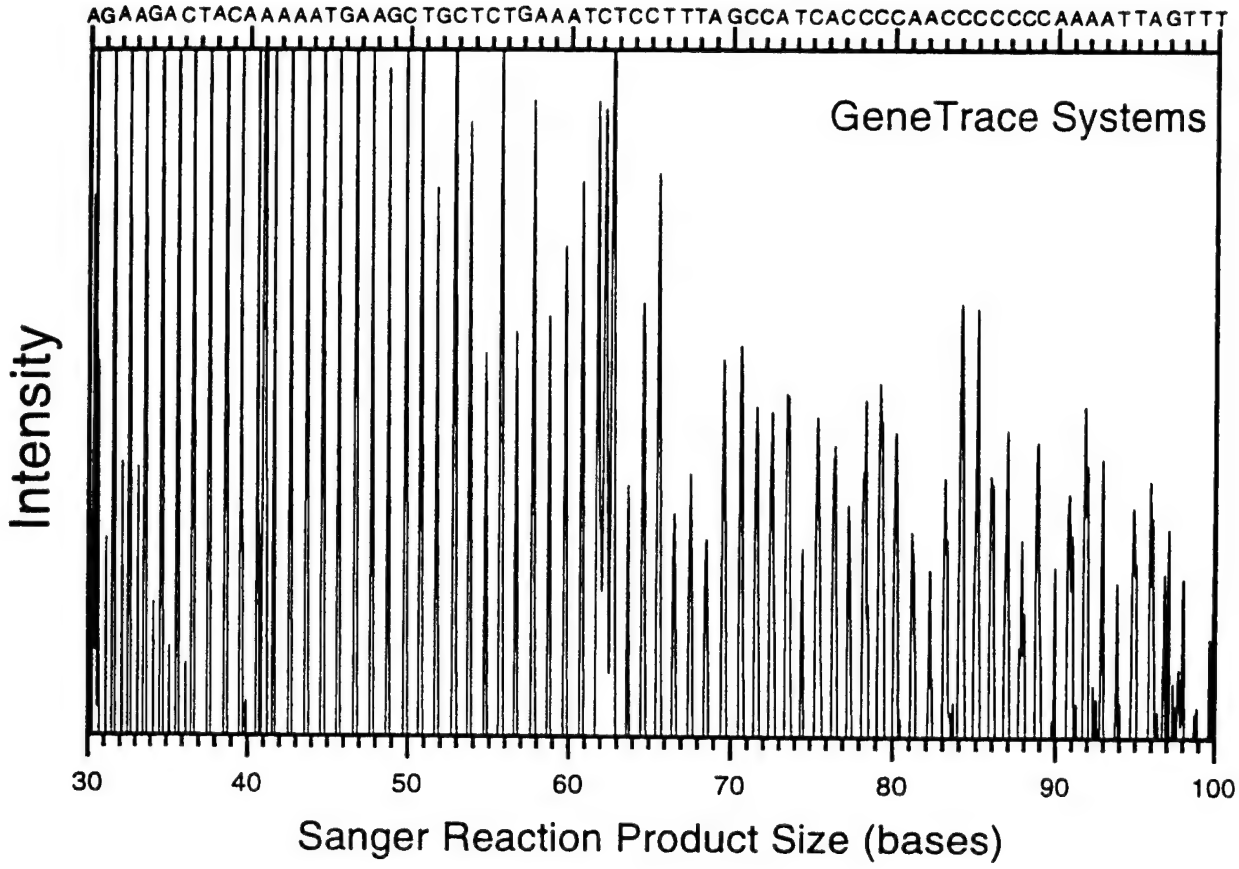


Figure 1. Time-of-flight mass spectrum of a DNA sequence from a subpicomolar amount of cDNA template using a Sanger protocol where all four dideoxys are copresent and the mass accuracy is used to call the sequence from the measured mass difference between peaks.

Another important class of assays involves sizing alleles of microsatellites (short nucleotide repeats); see Fig. 2. Principal applications include large scale genotyping for mapping disease-related genes[6], use as indicators of genetic instability associated with cancer[7], inherited genetic disorders such as fragile X syndrome and myotonic dystrophy[8], and pathogen/human/animal identification. DNA sizing can also be used for mutation screening. The high mass accuracy inherent to this type of analysis can be used to determine deviations from wild type in the size of an amplicon or fragments of an amplicon.

A further class of DNA assay monitors single nucleotide polymorphisms (SNPs), also known as biallelic markers. These markers can be used to identify subspecies of pathogens as well as general genotyping for all species. In these assays, a primer is extended by a single base by polymerase addition of a dideoxynucleoside triphosphate without the presence of deoxynucleoside triphosphates. These SNP samples can be multiplexed and heterozygotes detected. The base addition is determined by the mass difference between the unextended and extended primer using the high mass accuracy available (Fig. 3).

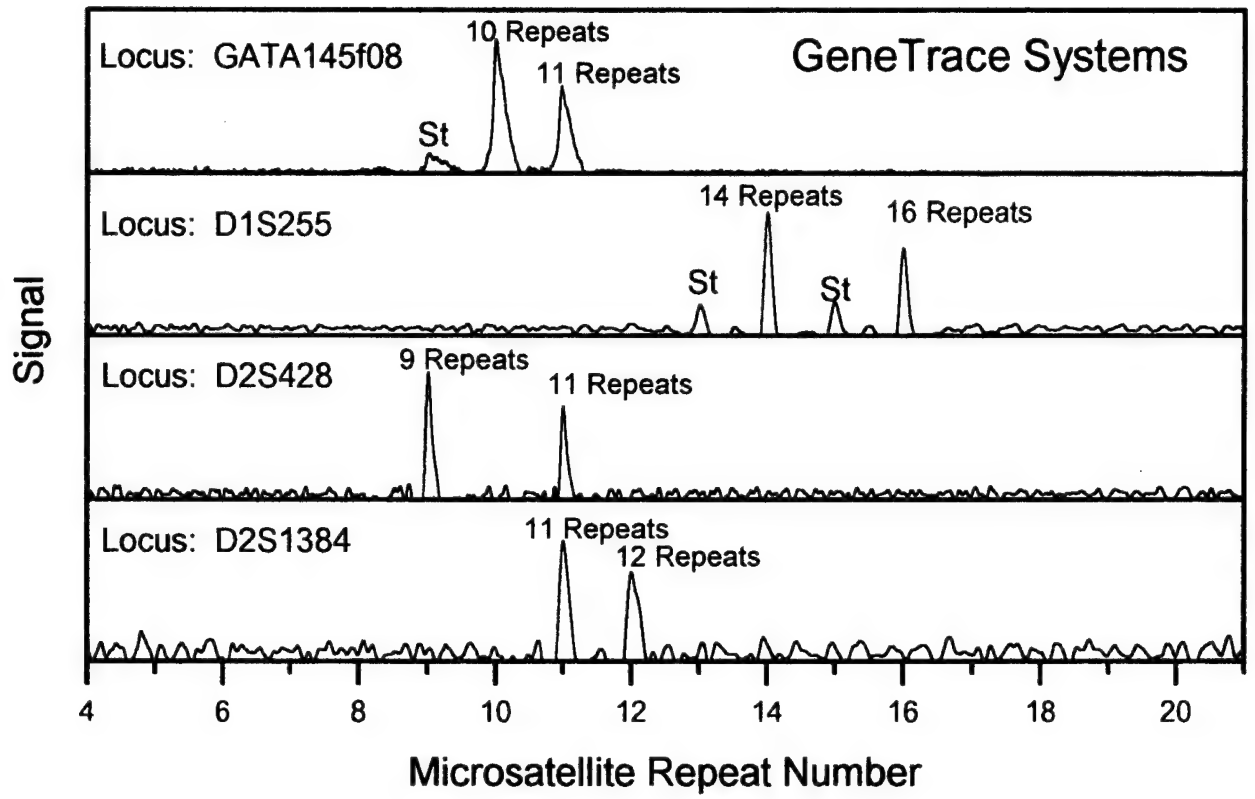


Figure 2. Time-of-flight mass spectrum of four different microsatellite loci from human genomic DNA, PCR amplified. 'St' = stutter (or slippage) products from the amplification process.

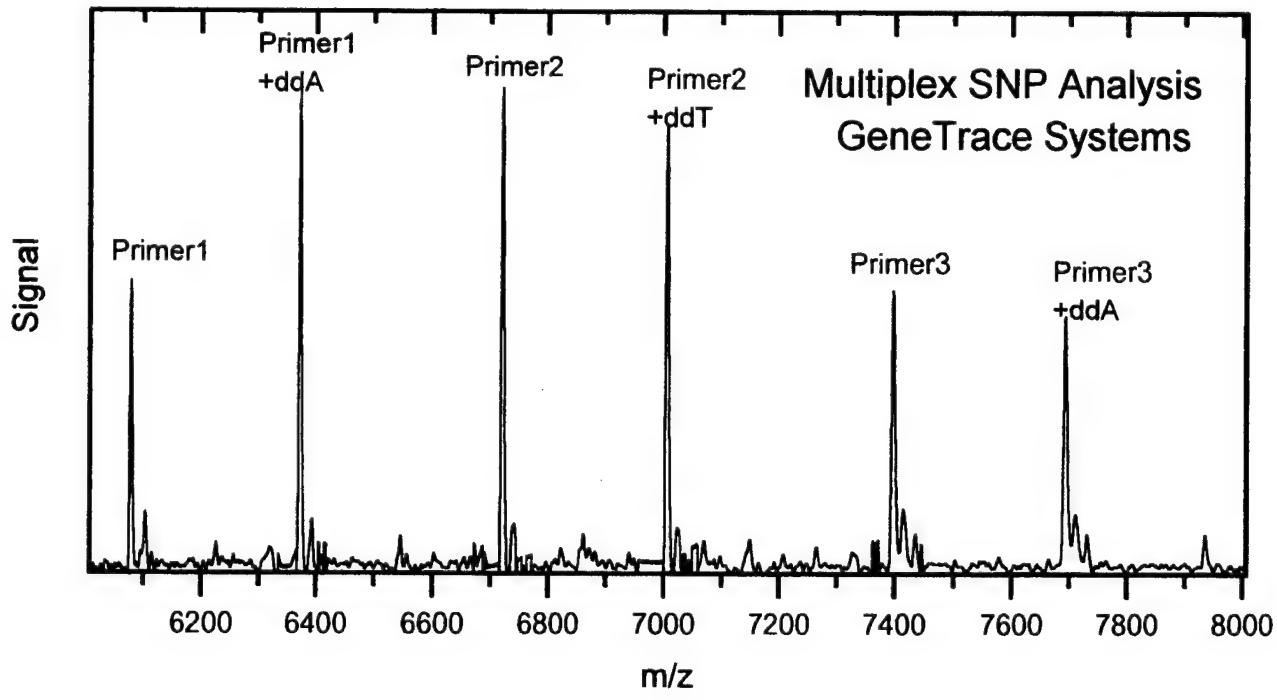


Figure 3. Time-of-flight mass spectrum of three different (multiplexed) SNP loci showing the (polymorphic) base present. Heterozygotes can also be detected.

2.2 NEW MATRICES FOR DNA ANALYSIS

We have developed a new matrix approach based on volatile (at room temperature) crystalline solids.[9] Two molecules, 8-hydroxyquinoline (8HQ) and 4-nitrophenol (4NP), are particularly effective as MALDI matrices for DNA oligomers. A liquid-nitrogen cooled sample stage maintains a low vapor pressure of the matrix during analysis in the time-of-flight mass spectrometer previously described.[10,11]

8HQ is an effective matrix for high-resolution analysis of DNA oligomers of less than approximately 100 nucleotides (30kDa). The 0.1 M 8HQ matrix solution was prepared in 1:1:2 (volume) acetone:butanone:water solution containing 25 mM ammonium citrate and 10 mM CDTA. 8HQ is commonly used to chelate trace amounts of metal ions, especially copper, but the addition of CDTA effectively suppressed copper adducts in the mass spectrum. Spectra of oligonucleotides in this matrix typically have a low background ion signal and high signal-to-noise levels. Figure 4 illustrates the mass resolution attainable for single-stranded oligonucleotides of about 27 kDa. DNA oligomers 90 and 91 nucleotides in length have a mass resolution (m/Δm) of 625 and 700, respectively, at full width at half height. The high sensitivity and resolution of mass spectra obtained from this matrix make it potentially useful for DNA sequencing applications.

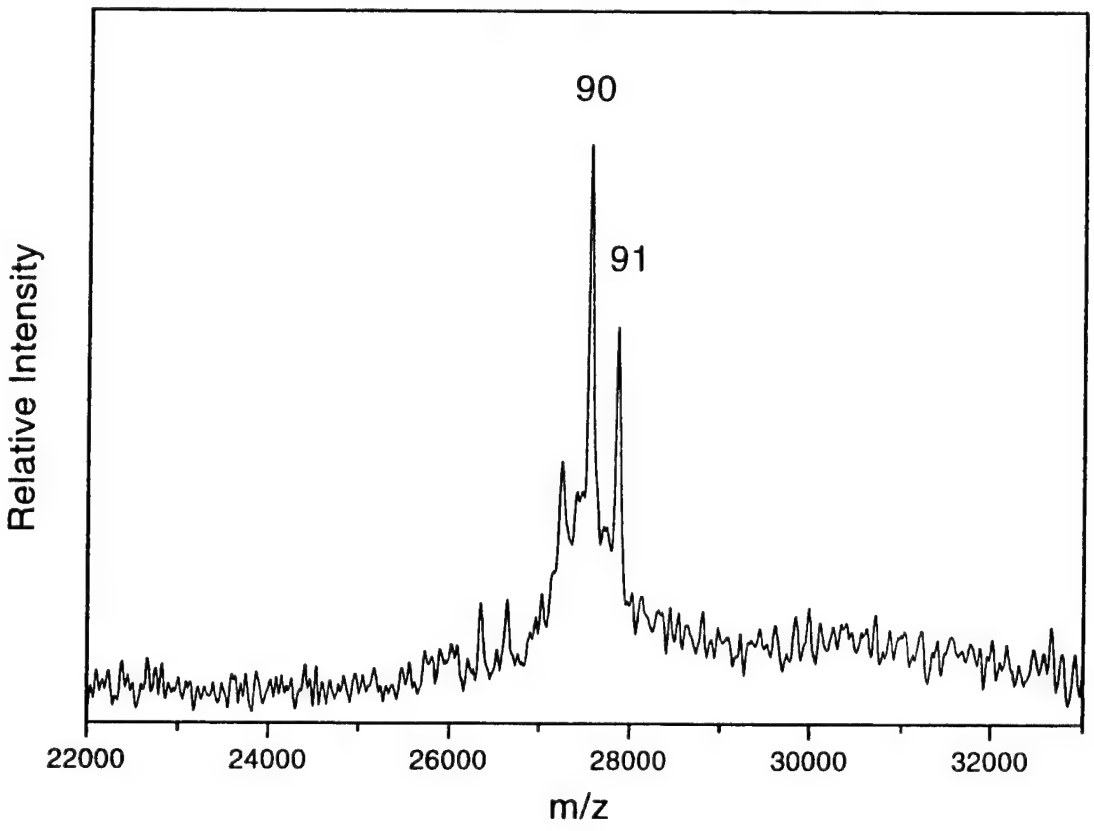


Figure 4. Positive ion mass spectrum of a mixture of single-stranded DNA oligomers 90 and 91 nucleotides in length obtained using 8-hydroxyquinoline as the matrix on a cooled silicon substrate. The spectrum is the sum of 200 laser pulses at 355 nm. The oligonucleotide sample was obtained from polymerase chain reaction (PCR) amplification of a short tandem repeat sequence at the human TH01 (tyrosine hydroxylase gene) locus.

4NP is especially effective for sensitive detection of higher mass (greater than approximately 30 kDa) DNA oligomers. The preparative solution for this matrix was 0.5 M 4NP in 1:1 (volume) methanol:water containing 50 mM diammonium citrate. The mass spectrum in Figure 5 shows the highest mass oligonucleotide detected. The parent peak at approximately 315 kDa corresponds to an estimated number of nucleotides exceeding 1,000. The width of the peak originates in part from the mass difference of the two complementary DNA strands (denatured during analysis) and partly from adduct formation as well as fragmentation. Because of the sensitive detection of long PCR products in this matrix, it is useful for DNA sizing applications.

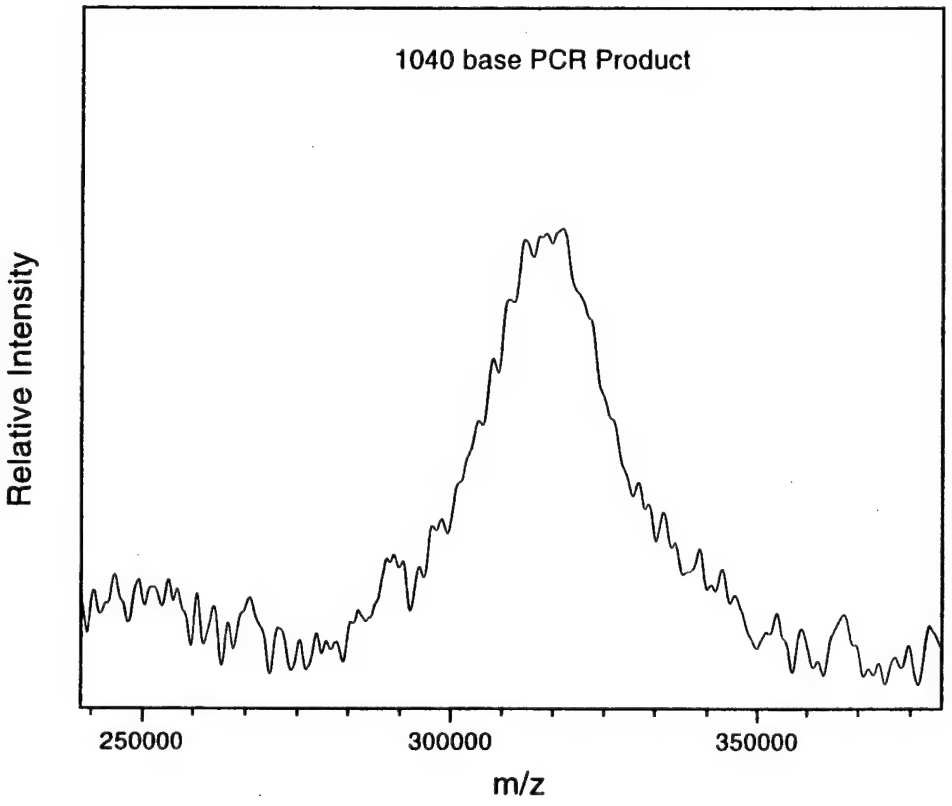


Figure 5. Positive ion mass spectrum of approximately 10 pmol of a double-stranded PCR product greater than approximately 1000 nucleotides in length using a 4-nitrophenol matrix on a cooled silicon substrate. The spectrum is the sum of 1000 laser pulses at 355 nm. The sample was derived from a cDNA insert in a plasmid vector.

3. CONCLUSIONS

Time-of-flight mass spectrometry of DNA can be expected to play a role in pathogen detection as well as pharmaceutical research and development and clinical diagnostics. Rapid and flexible DNA analysis can provide a powerful way of detecting low abundances of infectious pathogens, with the help of amplification methods such as PCR. The role of this technology in pharmaceutical research will be broad and powerful allowing the analysis of gene expression and gene mutation patterns, gene discovery and genetic mapping, as well as analysis of gene products. For clinical diagnostics, tests for the existence of

genetic defects involving base changes and nucleotide-repeat instabilities are foreseen. A key practical aspect is that such medical tests must be inexpensive as well as robust and rapid. The high degree of automation along with the inherent speed of this DNA analysis method suggests a place for this technology in the clinical setting in the next few years.

4. ACKNOWLEDGMENTS

This work was supported in part by the U.S. Department of Commerce, National Institute of Standards and Technology Advanced Technology Program, Cooperative Agreement No. 70NANB5H1029, and by the U.S. National Institutes of Health, National Human Genome Research Institute, Grant No. R01-HG00174.

5. REFERENCES

1. Karas, M. and Hillenkamp, F. *Anal. Chem.* **60**, 2299-2301 (1988).
2. Karas, M., et al. *Anal. Chimica Acta* **241**, 175-185 (1990).
3. Huth-Fehre, T., Gosine, J.N., Wu, K.J., and Becker, C.H. *Rapid Commun. Mass Spectrom.* **6**, 209-213 (1992).
4. Wu, K.J., Steding, A., and Becker, C.H. *Rapid Commun. Mass Spectrom.* **7**, 142-146 (1993).
5. Sanger, F., Nicklen, S., and Coulson, A.R. *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467 (1977).
6. Hall, J.M., LeDuc, C.A., Watson, A.R., and Roter, A.H. *Genome Research* **6**, 781-790 (1996).
7. Mao, L. et al. *Science* **271**, 659-662 (1996).
8. Genetic Disorders and the Fetus. (3rd edition, ed. Milunsky, A.), (The Johns Hopkins University Press, Baltimore and London, 1992).
9. Lin., H., Hunter, J.M., and Becker, C.H. To be published (1997).
10. Hunter, J. M.; Lin, H.; Becker, C. H. *Anal. Chem.* Accepted for publication (1997).
11. Wu, K. J.; Shaler, T. A.; Becker, C. H. *Anal. Chem.* **66**, 1637 (1994).

THE CRITICAL ROLE OF MASS SPECTROMETRY IN BIOTECHNOLOGY: GENE IDENTIFICATION AT THE PROTEIN LEVEL

Surinder Kaur, Dazhi Tang, Scott Chamberlain and Frank Masiarz
Protein Structure Analysis Group, Chiron Corporation
Emeryville, CA 94608-2916, USA

ABSTRACT

Current biological mass spectrometry methods are powerful tools which play a critical role in biotechnology by providing detailed structural information for recombinant proteins. The recent, explosive growth of sequence databases, together with advances in technology, offer the potential to revolutionize the role of mass spectrometry in biotechnology by allowing the rapid identification of cellular proteins. A number of research groups have developed mass spectrometry and database searching approaches to identify proteins isolated by gel electrophoresis. We show how we have used these approaches to identify an unknown, endogenous protein present in a preparation of a vaccine antigen expressed in CHO cells.

INTRODUCTION

Biological mass spectrometry (MS) methods, such as electrospray ionization MS (ES MS; Fenn *et al.*, 1989), nano-electrospray ionization (nanoES; Wilm *et al.*, 1996), tandem MS (MS/MS; Hunt *et al.*, 1986; Biemann and Scoble, 1987) and matrix-assisted laser desorption ionization time-of-flight MS (MALDI-TOF MS; Hillenkamp *et al.*, 1991), are powerful tools which play a critical role in a range of areas in biotechnology. Crucial information can be obtained rapidly for the discovery, development and structural characterization of therapeutic agents, vaccine antigens and diagnostic antigens.

For therapeutic proteins or vaccine antigens, the identification of post-translational modifications is important, as the modifications may be necessary for protein function. Post-translational modifications are also important since they can influence the biological half-life of a therapeutic protein. MS is a powerful way to characterize protein post-translational modifications in therapeutic proteins and antigens (Hemling *et al.*, 1995, Norais *et al.*, 1996a, 1996b). In small molecule drug discovery, a promising new MS strategy involves the combination of affinity selection and MS for identifying lead compounds in combinatorial libraries which interact with specific protein targets (Kaur *et al.*, 1997).

In gene discovery, the recent growth of sequence databases promises to revolutionize the role of MS in biotechnology by allowing the rapid identification of cellular proteins isolated by 1-D or 2-D gel electrophoresis (Shevchenko *et al.*, 1996; Mann, 1996; Yates *et al.*, 1996). In the proteome analysis approach, all the proteins in cells of interest, present in femtomole amounts, can be identified. The identification of differentially expressed proteins offers a powerful method to study disease states and therapeutic interventions at the protein level. In a more targeted approach, the biological function of gene products can be studied by the identification of multiprotein complexes (Lamond and Mann, 1997). There are two major MS approaches for protein identification. The latest high throughput approach identifies proteins by comparing the accurate molecular masses of peptides (<30 ppm) obtained by delayed-extraction MALDI reflectron TOF (Hillenkamp *et al.*, 1991, Brown and Lennon, 1995) directly with the theoretical peptide molecular masses for every protein sequence and open reading frame in the database (Jensen *et al.*, 1996). Alternatively, peptide fragment ions, generated during tandem MS or post-source decay MALDI-TOF MS, in conjunction with the intact peptide molecular mass, can be used for database searching (Shevchenko *et al.*, 1996; Mann, 1996; Yates *et al.*, 1996). New MS technologies, using hybrid instruments, promise high throughput identification of novel proteins where database sequences are not available (Shevchenko *et al.*, 1997). Here we show how we have used delayed-extraction MALDI reflectron TOF MS and nanoES MS,

with database searching, to identify an endogenous unknown protein which was co-purified with a vaccine antigen expressed in CHO cells.

RESULTS AND DISCUSSION

Figure 1 shows the 1-D polyacrylamide gel electrophoresis (PAGE, silver-stain) results of a glycosylated vaccine antigen preparation from CHO cells, where an unknown endogenous protein was co-purified. The broad lower band corresponds to the antigen and the sharp upper band corresponds to an unknown 70 kDa protein. The aim of the work was to identify the unknown protein. Figure 1 also shows the experimental scheme used for protein isolation and identification (Shevchenko *et al.*, 1996) The gel band corresponding to the endogenous protein was excised and the unknown protein was digested in-gel with trypsin. The tryptic peptides were extracted and the molecular ion masses were obtained by MALDI reflectron TOF MS using a Bruker REFLEX instrument (Bruker Analytical Systems, Inc., Billerica, MA) with a delayed-extraction ion source.

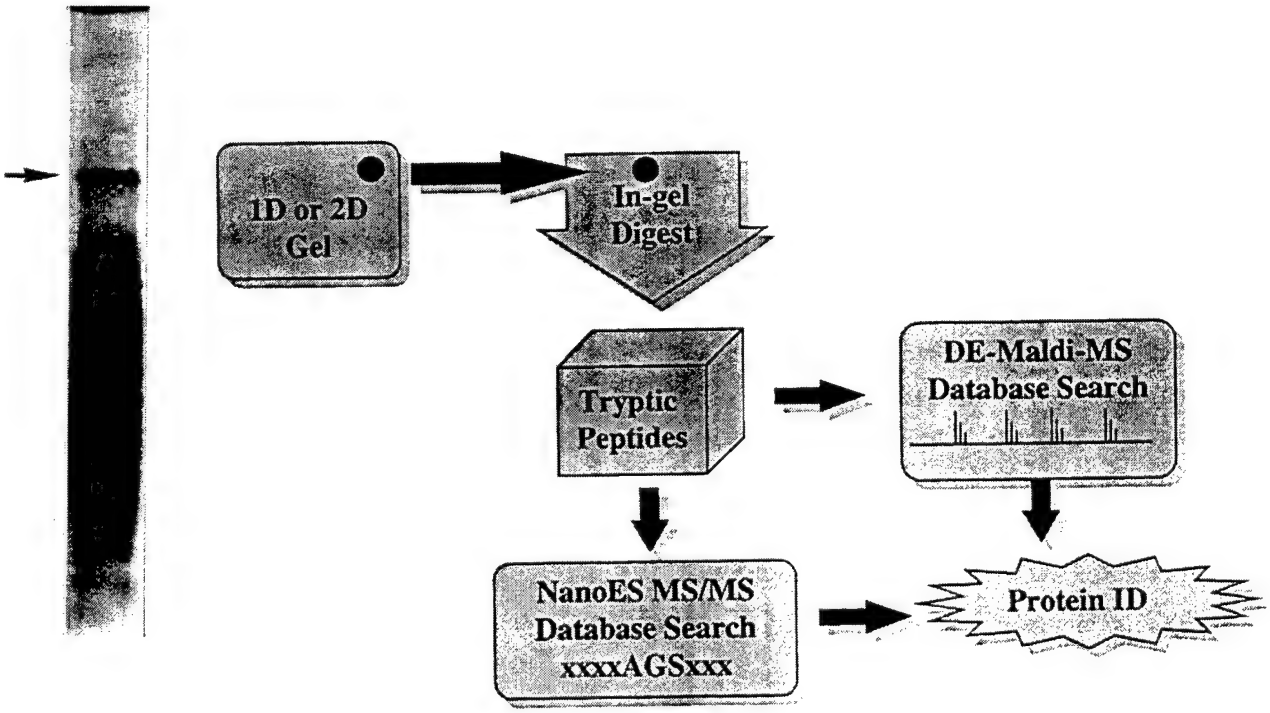


Figure 1. 1-D PAGE (silver stain) of a vaccine antigen preparation from CHO cells containing an unknown endogenous protein and the experimental scheme used for protein identification (Shevchenko *et al.*, 1996). → Unknown protein.

Figure 2 shows the MALDI-TOF MS mass spectrum of the peptide mixture obtained from the 70 kDa gel band corresponding to the unknown protein. A resolution (FWHM) of approximately 4000 or greater was obtained in the mass/charge (*m/z*) range 1000-2500. The mass spectrum was calibrated internally using a trypsin autolysis peptide (*m/z* 2163.06) and a matrix-related ion (*m/z* 568.13)

present in the spectrum. The inset shows isotopically resolved peptide molecular ions with the ^{12}C isotope having *m/z* 2149.02. The monoisotopic molecular ion masses of the unknown peptides were used to search a non-redundant protein database using PeptideSearch software (Mann, 1994).

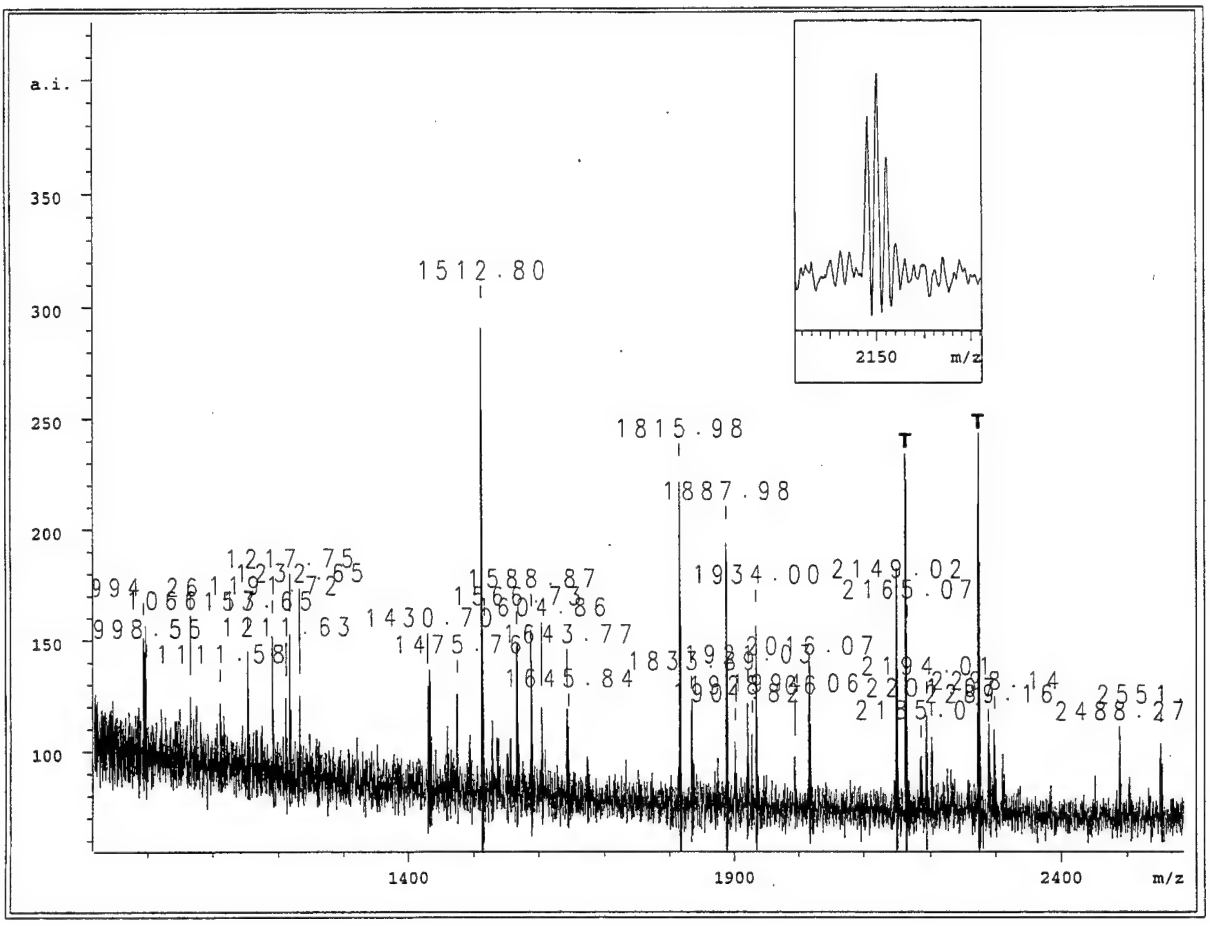
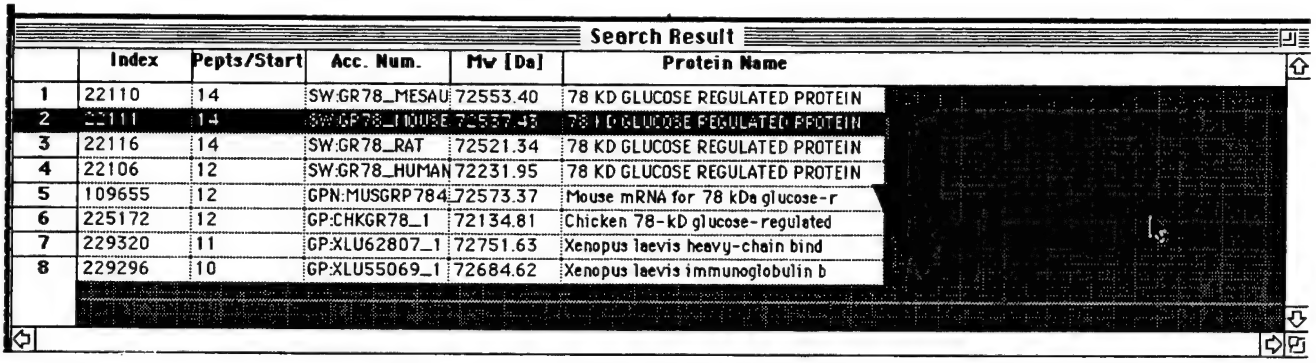


Figure 2. Delayed-extraction MALDI reflectron TOF mass spectrum of tryptic peptides from gel-isolated unknown protein. **T**: Trypsin autolysis peptides used as internal calibrants.

Table 1 shows the results of the protein database search which identified the unknown protein as a glucose regulated protein (GR78). Positive matches were found for the same protein from a variety of species, including rodents, which is consistent with a CHO cell source. The protein GR78 is a member of the stress family of proteins and is a "molecular chaperone" involved in the transport of folded proteins (Flynn *et al.*, 1991). Table 2 shows the monoisotopic molecular ion masses measured by delayed-extraction MALDI-TOF and compares them with the theoretical predicted molecular ion masses for 16 of the GR78 tryptic peptides. All except one of the low abundance peptides showed experimental molecular ion masses which were within 2-30 ppm of the predicted theoretical molecular ion masses using the two point calibration above. The 16 peptides with positive matches represent 31% of the amino acid sequence.



Search Result

	Index	Pepts/Start	Acc. Num.	Mw [Da]	Protein Name
1	22110	14	SW:GR78_MESAU	72553.40	78 KD GLUCOSE REGULATED PROTEIN
2	22111	14	SW:GR78_MOUSE	72337.43	78 KD GLUCOSE REGULATED PROTEIN
3	22116	14	SW:GR78_RAT	72521.34	78 KD GLUCOSE REGULATED PROTEIN
4	22106	12	SW:GR78_HUMAN	72231.95	78 KD GLUCOSE REGULATED PROTEIN
5	109655	12	GPN:MUSGRP784	72573.37	Mouse mRNA for 78 kDa glucose-r
6	225172	12	GP:CHKGR78_1	72134.81	Chicken 78-kD glucose-regulated
7	229320	11	GP:XLU62807_1	72751.63	Xenopus laevis heavy-chain bind
8	229296	10	GP:XLU55069_1	72684.62	Xenopus laevis immunoglobulin b

Table 1. Positive matches from database search using PeptideSearch software

	Measured	Calculated	(+ H^+)	Diff.	Residues	Sequence
	1153.650	1153.618	(mono)	0.032	298-307	(K)RALSSQHQAR(I)
	1191.720	1191.637	(mono)	0.083	466-475	(K)VYEGERPLTK(D)
	1430.700	1430.691	(mono)	0.009	103-114	(R)TWNDPSVQQDIK(F)
	1512.800	1512.751	(mono)	0.049	326-337	(R)AKFEELNMDLFR(S)
*	1566.730	1566.780	(mono)	-0.050	62-75	(R)ITPSYVAFTPEGER(L)
	1588.870	1588.854	(mono)	0.016	354-368	(K)KSDIDEIVLVGGSTR(I)
	1604.860	1604.864	(mono)	-0.004	125-139	(K)TKPYIQVDIGGGQTK(T)
	1645.840	1645.869	(mono)	-0.029	183-198	(R)QATKDAGTIAGLNVMR(I)
	1815.980	1815.996	(mono)	-0.016	199-215	(R)IINEPTAAAIAYGLDKR(E)
*	1833.890	1833.909	(mono)	-0.019	83-98	(K)NQLTSNPENTVFDAKR(L)
*	1887.980	1887.971	(mono)	0.009	166-182	(K)VTHAVVTVPAYFNDAQR(Q)
	1934.000	1934.013	(mono)	-0.013	476-493	(K)DNHLLGTFDLTGIPPAPR(G)
	1994.060	1994.100	(mono)	-0.040	291-307	(R)REVEKAKRALSSQHQAR(I)
	2016.070	2016.066	(mono)	0.004	165-182	(K)KVTHAVVTVPAYFNDAQR(Q)
	2149.020	2148.997	(mono)	0.023	308-325	(R)IEIESFFEGEDFSETLTR(A)
	2488.270	2488.314	(mono)	-0.044	346-368	(K)VLEDSDLKKSDIDEIVLVGGSTR(I)

Table 2. Measured and theoretical molecular ion masses of peptides from glucose regulated protein GR78 (SW:GR78_MOUSE). *Sequence tags obtained for peptides indicated.

Further structural information about the peptides was obtained by nanoES MS/MS sequencing (Wilm *et al.*, 1996). This technique involves fragmentation of a peptide by collision-induced dissociation (CID). Peptides in a mixture such as above can be individually analyzed in the mass spectrometer without prior chromatographic isolation. Partial interpretation of the fragments in the CID mass spectrum can be used to obtain a short sequence tag. The sequence tag, peptide molecular ion mass and related information can be used for database searching using PeptideSearch software (Mann, 1994). The desalted tryptic peptide mixture above was analyzed by nanoES MS/MS and sequence tags were obtained for the peptides indicated in Table 2. In each case, database searching showed a positive match for GR78. Figure 3 shows the nanoES MS/MS mass spectrum for the triply-

charged molecular ion for the peptide with *m/z* 1887.98, where the sequence tag PAY was identified. Database searching using *m/z* 1887.98 and the sequence tag PAY showed a positive match for GR78 and identified the sequence VTHAVVTVTVPAYFNDAQR in the protein. To confirm that all of the ions in the nanoES MS/MS mass spectrum corresponded to the proposed sequence, the theoretical fragments were calculated and are labeled in the spectrum (Figure 3).

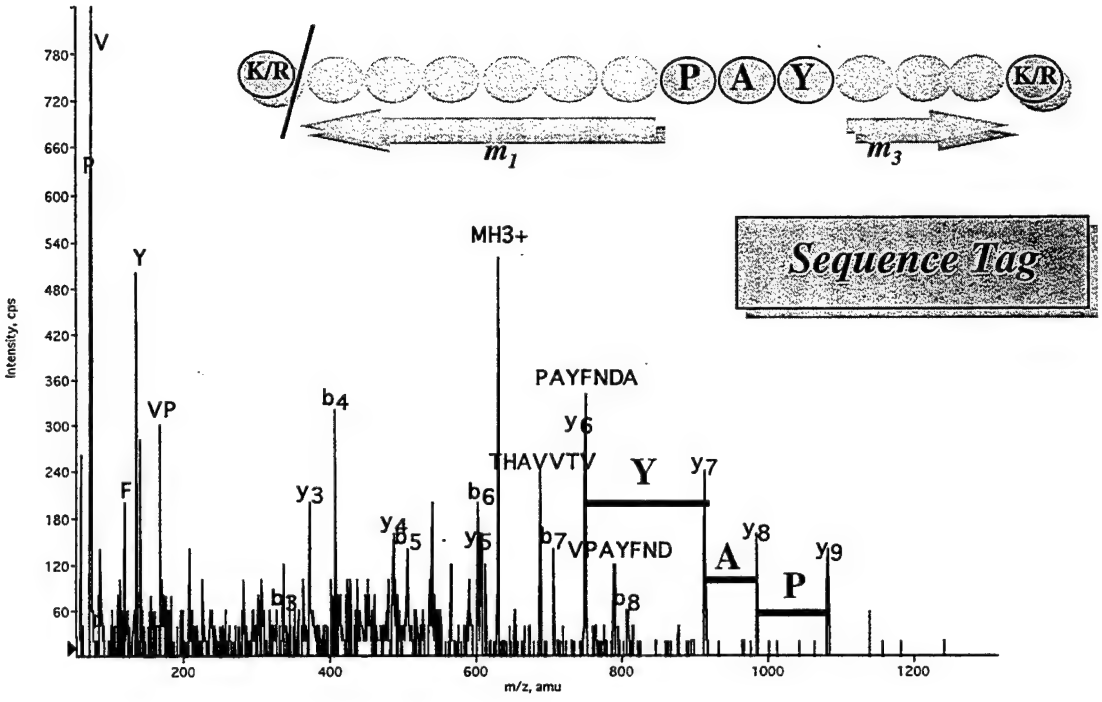


Figure 3. NanoES MS/MS mass spectrum of peptide with *m/z* 1887.98. Sequence tag PAY. The ions labeled in the spectrum correspond to the sequence VTHAVVTVTVPAYFNDAQR.

CONCLUSIONS

A mass spectrometry and database searching approach is a powerful method for identifying proteins isolated by PAGE. We have shown how we have used this approach to identify an unknown, endogenous protein in a preparation of a vaccine antigen as a "chaperone" protein GR78. Since the development of ES and MALDI, mass spectrometry has played a critical role in biotechnology by providing protein structural information, such as post-translational modifications, for recombinant proteins. The ability to identify femtomole levels of cellular proteins rapidly by MS and database searching, or *de novo* sequencing, is likely to have an equally significant impact.

ACKNOWLEDGMENTS

We are grateful to Yiu-Lian Fong and Larry Cousens of Chiron Corporation for the preparation of the protein sample.

REFERENCES

Biemann K., and Scoble, H. A. (1987) *Science,* **237**, 992-998.

Brown, R. S., and Lennon, J. J. (1995) *Anal. Chem.,* **67**, 1998.

Fenn, J. B., Mann M., Meng, C. K., Wong, S. F., and Whitehouse, C. M. (1989) *Science,* **246**, 64-71.

Flynn, G. C., Pohl, J., Flocco, M. T., and Rothman, J. E. (1991) *Nature*, **353**, 726.

Hemling, M. E., Mentzer, M. A., Capiau, C., and Carr, S. A. (1995) in *Mass Spectrometry in the Biological Sciences* (A. L. Burlingame and S. A. Carr, Eds.), Humana Press, pp. 307-331.

Hillenkamp, F., Karas, M., Beavis, R. C., and Chait, B. T. (1991) *Anal Chem.*, **63**, 1193A.

Hunt, D. F., Yates, J. R., III, Shabanowitz, J., Winston S., and Hauer C. R. (1986) *Proc. Natl. Acad. Sci. USA,* **83**, 6233-6237.

Kaur, S., McGuire, L., Tang, D., Dollinger, G., and Huebner, V. (1997) *J. Protein Chem.,* **16**, 505-511.

Jensen, O. N., Podtelejnikov, A., and Mann, M. (1996) *Rapid. Commun. Mass Spectrom.,* **10**, 1371-1378.

Lamond, A. I., and Mann, M. (1997) *Trends Cell Biol.* **7**, 139-142.

Mann, M. (1994) in *Microcharacterization of Proteins* (Kellner, R., Lottspeich, F. and Meyer, H. E., Eds.), pp. 223-245.

Mann, M. (1996) *TIBS*, **21**, 494-495.

Norais, N., Hall, J. A., Gross, L., Tang, D., Kaur, S., Chamberlain, S. H., Burke R. L., and Marcus, F. (1996a) *J. Virol.,* **70**, 5716-5719.

Norais, N., Tang, D., Kaur, S., Chamberlain, S. H., Masiarz, F. R., Burke R. L., and Marcus, F. (1996b) *J. Virol.,* **70**, 7379-7387.

Shevchenko, A., Jensen, O. N., Podtelejnikov, A. V., Sagliocco, F., Wilm, M., Vorm, O., Mortensen, P., Shevchenko A., Boucherie, H., and Mann, M. (1996) *Proc. Natl. Acad. Sci. USA*, **93**, 14440-14445.

Shevchenko, A., Chernushevich, I., Ens, W., Standing, K. G., Thomson, B., Wilm, M., and Mann, M. (1997) *Rapid Commun. Mass Spectrom.,* **11**, 1015-1024.

Wilm, M., Shevchenko, A., Houthaeve, T., Breit, S., Schweigerer, L., Fotsis, T., and Mann, M (1996) *Nature (London),* **379**, 466-469.

Jensen, O. N., Podtelejnikov, A., V., and Mann, M. (1996) *Rapid Commun. Mass Spectrom.,* **10,** 371-378.

Yates, J. R., III, McCormack, A. L., and Eng, J. (1996) *Anal. Chem.,* **68**, 534A-540A.

MARKERS FOR IDENTIFICATION/DETECTION OF BIOLOGICAL AGENTS ASSESSED USING MASS SPECTROMETRY

Alvin Fox
Department of Microbiology & Immunology
University of South Carolina
School of Medicine
Columbia, SC 29208

Bacterial pathogens are readily identified by their genetic constituents using the polymerase chain amplification reaction (PCR). PCR has great specificity allowing biodetection in complex environmental samples (e.g. aerosols). These PCR products are generally identified by their molecular weight (M. W) as determined by electrophoresis; this is extremely time consuming. Alternatively M.W. can be determined rapidly by electrospray mass spectrometry. Tandem mass spectrometry (MS/MS) potentially allows even greater discriminating power since a "fingerprint" (product spectra) of oligonucleotide fragments of the PCR product may be generated in real-time.

INTRODUCTION

The aim of biodetection is to rapidly detect microbes during an airborne attack. Once bacteria are concentrated from air, biodetection strategies include plans to exploit chemical differences among microbes using mass spectrometric-based approaches. Virulence factors of bacteria include structural proteins (e.g. anthrax toxin) or synthetic enzymes involved in synthesis of non-protein structures (e.g. the poly D-glutamic acid capsule of *B. anthracis*). In either case, specific proteins makers for microbial identification are encoded at the genetic level by corresponding nucleotide sequences. Such marker sequences are also highly specific for each key biological agent (e.g. *Brucella melitensis, Bacillus anthracis*, *Yersinia pestis* and *Francisella tularensis*). A region of more general applicability is the 16S/23S ribosomal RNA interspace region (Wunschel, *et al.*, 1994). This region is present in all bacterial genomes. However, the sequence and nucleotide length of the spacer region differs among bacteria in a species specific fashion. Unfortunately, many of these protein and nucleic acid markers are present at low levels in the bacterial cell. Furthermore, these bacterial species constitute only minor components of complex environmental samples. Thus isolation of sufficient amounts of markers suitable for analysis is currently problematic.

Alternatively, using polymerase chain amplification (PCR) millions of copies of genetic sequences can be made rapidly by enzymatic amplification. The detection of these sequences conventionally involves gel electrophoresis which is extremely time consuming. However, using electrospray mass spectrometry (ES-MS) characterization of PCR products could potentially be achieved in a matter of few seconds (Muddiman et. al., 1996, Wunschel, *et al.*, 1996; Naito, *et al.*, 1995; Tsuneyoshi, *et al.*, 1997). Comparable results are achieved by in MS analysis of PCR products using matrix assisted laser desorption (MALDI) MS (Doktycz, *et al.*, Liu, *et al.*, 1995; 1995; Ross, 1997); however it is more difficult to adapt this approach to routine MS/MS analysis. MS merely provides a measure of M.W. Alternatively, tandem mass spectrometry (MS/MS) potentially allows even greater discriminating power since a "fingerprint" (product spectra) of oligonucleotide fragments can be generated in real-time (McLuckey, *et al.*, 1992; Ni, *et al.*, 1996, Little, *et al.*, 1995). The purpose of the our research is to develop a general scheme to rapidly analyze PCR products for biodetection using MS and MS/MS.

Synthetic oligonucleotides (over 100 mers in size) are readily ionized by electrospray mass spectrometry with high sensitivity. Due to the presence of phosphate groups between each nucleotide subunit, analysis is performed in the negative ion mode. It is well established that these highly negatively charged DNA

molecules tend to form adducts with positively charged mono and divalent cations. These adducts can complicate the spectrum increasing the complexity of mass spectral interpretation. Thus PCR inherently involves some additional chemical clean-up problems. High concentrations of low molecular weight nucleic acids, "primers" (around 20 nucleotides in length) as well the nucleotide monomeric building blocks are integral to amplification of the target genetic information. These low molecular weight nucleotides tend to ionize much more efficiently than the higher molecular weight PCR products. Thus a scheme for analysis of PCR products by ES MS or ESI MS/MS must include removal of cations, low molecular weight nucleotide products as well as other components of the amplification mixture (Muddiman, *et al.*, 1996).

RESULTS

Determination of single base changes by FTICR MS analysis. An MS procedure for rapid and precise analysis of PCR products using has been developed in a collaboration between the University of South Carolina (USC) and Pacific Northwest National Laboratory (PNNL). The initial work involved the use of a high resolution FTICR instrument which allowed rapid determination of the MW of each of the strands of a double stranded PCR product (Muddiman, *et al.*, 1996). The accuracy of the MW determination allowed recognition of single base substitutions, deletions and additions .

A G (guanine, MW: 151) to C (cytosine, MW: 111) switch produces a molecular weight change of ~40 Da and an A (adenine, MW: 135) to T (thymine. MW:126) represents a change of ~9 Da. The most important aspect of MS analysis for evaluation of PCR products is the exquisite accuracy of the MW determination. Not only can changes be immediately detected but for single base changes the difference in MW is specific for only one base, and can be confirmed with the complementary change in the opposite strand. Thus, analysis of both strands produces information allowing any base composition change to be assignable. A second important feature of MS analysis is the speed at which results can be obtained. Minimal sample preparation is required, taking at most 10-15 min and rapid MS analysis can be achieved in a few seconds, allowing a large number of sequences to be rapidly determined. The sample preparation and analysis are easily adapted to automation for even faster on-line analysis and processing of large numbers of samples (Muddiman, *et al.*, 1996)

FTICR MS data agrees with Sanger sequence determined by electrophoresis. The MW of the PCR product as determined by MS analysis were in excellent agreement with predicted nucleic acid sequence (determined by conventional electrophoresis using Sanger sequencing ladders). When discrepancies between MW determination by MS and reported nucleic acid sequence were investigated by traditional sequencing the base substitutions corresponding to the mass differences were confirmed and their exact location identified. For example, the mass measurements of single strands of W23 were not in agreement with those calculated based on the sequence of another strain, 168 reported in the literature. The measured mass for the sense strand was determined to be ca. 40 Da less than the predicted mass and the antisense strand ca. 40 Da higher than the predicted mass. This mass discrepancy between the two strands clearly suggested a C to G switch between the two strands, resulting in a 40.03 Da addition to one strand and a corresponding 40.03 subtraction in mass in the other. The PCR product from strain W23 was subsequently cloned and sequenced by our group. As predicted from the MS data there was indeed a G to C switch. The switch was at position 49. (Krahmer, *et al.*, 1997); see Figure 1 below:

Sequence of PCR product from *B. subtilis* strain W23

GTGTTCTTTG[10] AAAACTAGAT[20] AACAGTAGAC[30] ATCACATTCA[40] ATTAGTAAGA[50]
CAAGATATCA[60] CATAGTGATT[70] CTTTTTAACG[80] GTTAAGTTAG[90] AAAGGGCGCA[100]
CGGTGGATGC[110] CTTG[114]

Sequence of PCR product from *B. subtilis* strain 168 (operon *rrnB*)

GTGTTCTTTG[10] AAAACTAGAT[20] AACAGTAGAC[30] ATCACATTCA[40] ATTAGTAACA[50]
CAAGATATCA[60] CATAGTGATT[70] CTTTTTAACG[80] GTTAAGTTAG[90] AAAGGGCGCA[100]
CGGTGGATGC[110] CTTG[114]

From this work, a sequence change was determined by observation of the MW changes of PCR products from two related strains of bacilli (168 and W23) by MS and conventional sequencing was used to verify the sequence change.

PCR products can be analyzed with a quadrupole instrument. The 7 Tesla FTICR instrument used in studies described above was developed in Dr. Smith's laboratory and it's capability exceeds many commercial instruments. ESI FTICR MS analysis of PCR products has proven the ability of MS in detection of single base changes within moderately sized PCR products. These state-of-the-art instruments are, however, not currently amenable to routine analysis with large throughput. It is important to demonstrate that PCR products can be analyzed on commercial mass spectrometers of lower cost. PCR products were successfully analyzed with a Quattro 1 triple quadrupole instrument. As anticipated, quadrupole MS does not currently have the resolution of FTICR MS (Krahmer, *et al.*, 1997). However, each pair of peaks derived from the two single DNA strands respectively were are well resolved. The resolution achieved would be adequate to detect base deletions or additions.

Quadrupole and ion trap MS analyses give readily comparable information. Ion trap MS instruments (e.g. LCQ) are modest in cost compared to triple quadrupole instruments. Using the commercial quadrupole instrument, synthetic oligonucleotides (30 mers) which differed by one base (C to G substitution) had a measured mass difference of 41.7 mass units which agreed with the predicted mass difference of 40. The MS analysis of oligonucleotides with the quadrupole instrument generated molecules of the same mass/charge ratio as observed with an ion trap. Although several additional fragments of higher mass/charge were observed in mass spectra generated with the ion trap.

MS/MS differentiation of nucleotide sequences using an ion trap. When used for MS/MS fingerprinting (product spectra), the LCQ MS/MS ion trap gave superior results in terms of sensitivity over the Micromass triple quadrupole instrument. In collaborative studies, between Mark Krahmer, Karen Fox, and Alvin Fox (USC) with Mona Shahgholi and John Callahan, Naval Research Laboratories (NRL) MS/MS ion trap (LCQ) analysis differentiated two synthetic 30 mers that differed only by a C to G substitution produced distinct fingerprints. For example, the higher MW parent ions (mass/charge of 918.1 and 921.7 respectively) generated a dominant product ion (mass/charge of 903.5 and m/z 907.5 respectively). It is noteworthy there are several other abundant ions that readily distinguished the two fingerprints. These results are encouraging and suggest that MS/MS analysis of PCR products may be equally successful.

CONCLUSIONS

The use of PCR-MS and PCR-MS/MS present exciting possibilities for specific identification of microbial species in complex matrices. It appears likely that this can be achieved with extreme rapidity, simplicity and sensitivity. The key advantage of analysis of PCR products in contrast to any other chemical component of bacteria is that the PCR reaction generates large amounts of the sequence of interest. Thus PCR products can be readily detected without the need for chromatographic or electrophoretic isolation of the material from other components in complex environmental samples. PCR and MS are a powerful combination since rapid characterization of PCR products by M.W. (analogous to gel electrophoresis) is possible. MS/MS fingerprints (product spectra) represent an even more powerful approach, since (like restriction mapping) they reflect the underlying sequence of the PCR product. The presence of specific bacterial species in complex environmental matrices would be definitively identified by their MS/MS fingerprints.

REFERENCES

1. Doktycz M., Hurst G., Habibi Goudarzi S., McLuckey S., Tang K., Chen C., Uziel M., Jacobson K.B., Woychik R. and Buchanan M. Analysis of polymerase chain reaction-amplified DNA products by mass spectrometry using laser desorption and electrospray: current status. Analyt. Biochem. 230: 205-214. 1995.
2. Krahmer M., Wunschel D., Fox. K., Fox A., Nagpal M., Kim K., Muddiman, D.C., Hofstaedler S. and Smith R. Genomic heterogeneity studied by ESI MS using both Fourier transform ion cyclotron resonance and triple quadrupole instruments. 45th ASMS Conference on Mass Spectrometry and Applied Topics. 1997.
3. Little P. and McLafferty F. Sequencing of 50-mer DNAs using electrospray tandem mass spectrometry and complimentary fragmentation methods. J. Am. Chem. Soc. 117: 6783-6784. 1995.
4. Liu Y.-H., Bai J., Zhu Y., Liang X., Siemieniak D., Venta P. and Lubman D. Rapid screening of genetic polymorphisms using buccal cell DNA with detection by matrix assisted desorption/ionization mass spectrometry. Rapid. Meth. Commun. Mass Spectrom. 9: 735-743. 1995.
5. McLuckey, S., Van Berkel G. J. and Glish G. L. Tandem mass spectrometry of small, multiply charged oligonucleotides. J. Am. Soc. Mass Spectrom. 3: 60-70. 1992.
6. Muddiman, D., Wunschel D, Liu, Pasa-Tolic L., Fox K., Fox A., Anderson G. and Smith R.D. Characterization of PCR products from bacilli using electrospray ionization FTICR mass spectrometry. Analyt. Chem. 68: 3705-3712. 1996.
7. Naito Y., Ishikawa K. and Koga Y., Tsuneyoshi T., Terunuma H. and Arakawa R. Molecular mass measurement of polymerase chain reaction products amplified from human blood DNA by electrospray ionization mass spectrometry. Rapid Meth. Commun. Mass Spectrom. 9: 1484-1486. 1995.
8. Ni, J., Pomerantz S., Rozenski J., Zhang, Y. and McCloskey J. Interpretation of oligonucleotide mass spectra for determination of sequence using electrospray ionization and tandem mass spectrometry. Analyt. Chem. 68: 1989- 1999. 1996.
9. Ross P. Oral Presentation. Joint services Workshop on Biological Mass Spectrometry, DoD, Baltimore, Maryland. July 1997
10. Tsuneyoshi T., Ishikawa K., Koga Y., Naito Y. , Baba S., Terunama H., Arakawa R. and Prockop D. Mass spectrometric gene diagnosis of one-base substitution from polymerase chain reaction amplified human DNA. Rapid Commun. Mass Spectrom. 11: 719-722. 1997
11. Wunschel D., Fox K, Black G.and Fox A. Discrimination among the *B. cereus* group, in comparison to *B. subtilis*, by structural carbohydrate profiles and ribosomal RNA spacer region PCR. System. Apply. Microbial. 17: 625-635. 1994.
12. Wunschel D.., Fox K., Fox. A., Bruce J., Muddiman D. and Smith R. Analysis of double stranded polymerase chain reaction products from the *Bacillus cereus* group by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. Rapid Commun. Mass Spectrom., 10: 29-35. 1996.

RAPID, AUTOMATABLE METHOD TO DETECT AND IDENTIFY THE INFECTIOUS POTENTIAL OF BW AGENTS USING SIGNATURE BIOMARKERS

D. C. White[1,2], R. S. Burkhalter[1], C. A. Flemming[1], K. T. Leung[1], K. W. Whitaker[3], J. S. Almeida[4], and S. J. Macnaughton[1,3]

[1]Center for Environmental Biotechnology, University of Tennessee, Knoxville, TN 37932-2575 USA, [2]Environmental Science Division, Oak Ridge National Laboratory, Oak Ridge, TN, [3]Microbial Insights, Inc., Rockford, TN, [4]University of Lisbon, Lisbon, Portugal

ABSTRACT: The rapid detection and identification of classical and genetically engineered biological warfare (BW) agents is crucial; whether the threat is a massive attack or a more subtle terrorist action. For infection, all agents including viruses require living cells which must have an intact polar lipid membrane. Analysis of signature lipid biomarkers (SLB) from these membranes provides a quantitative *in situ* method to detect BW agents (with the exception of prions and some virus particles). This method involves the extraction of membrane lipids from any number of matrices including aerosol filter retentates, mud, soils, aqueous filter retentates, vectors, and clinical specimens. Automated high-pressure, high-temperature solvent extraction yields SLB in minutes. SLB analysis involves fractionation, on-line derivatization, and separation and identification using GC/MS or HPLC/ES/MS. Other instrumentation combinations may be useful such as tandem mass spectrometry (MS/MS) for increased selectivity, greater structural information resulting from collisionally induced dissociation's (CID) and greater sensitivity resulting from the reduction of chemical noise; electrospray ionization (ES) for increased speed and sensitivity; HPLC for purification and additional structural conformation purposes; as well as the use of ion trap mass analyzers for field system portability along with MS^n capabilities. Since SLB's reflect both genotypic and phenotypic BW agent response, SLB's can provide an estimation of the infectious potential of specific agents in complex environmental matrices. The SLB technique has demonstrated the ability to detect the infectious potential of *Cryptosporidium parvum* at the sensitivity of a single oocyst using SIM/MS. The identification, pathogenicity, and drug-resistance of *Mycobacteria* spp. have been determined at sub-femtomolar sensitivities with SLB analysis. Examination of SLB's from aerosol filter retentates has also been utilize to provide quantitative detection and identification of unculturable, but viable and often infectious indoor air biocontaminants, as well as immune potentiating lipopolysaccharide-endotoxins and mycotoxins.

INTRODUCTION: Over the past 20 years, our laboratory has developed the signature membrane biomarker (SLB) method of analysis for the quantitative definition of the viable biomass, community composition, and nutritional status of microbiota isolated from a wide variety of environmental matrices including air, soil, wounds, and water [1]. In many cases, unique molecules specific to microorganisms that could be useful in detecting, identifying, and determining the infectious potential of life threatening microbial pathogens have been detected.

THE PROCESS:

Collection of the samples: A SLB based microbial testing system for airborne microorganisms based on the collection of particulates on glass fiber filters has been utilized to establish that the standard culturing

methods are woefully ineffective in determining the actual number of microbes in indoor air samples [2,3]. The SLB analysis system is quantitative, allows detection of lipopolysaccharide immune potentiators, and through examination of the phospholipid fatty acid (PLFA) profile provides clear evidence of metabolic stress [2,3]. Through this research it became clear that the lipid extraction approach provides an excellent method for the detection and identification of airborne biohazards. The major limiting factors in this technique are the speed of extraction and detection.

Extraction: The most time consuming step in the SLB analysis is the extraction of the lipids. The initial step classically involved a one-phase modified Bligh and Dyer solvent extraction for obtaining phospholipid fatty acid biomarkers used in community analysis. This method, however, is relatively labor intensive and slow, often taking up to 24 hours for the initial extraction.

Our laboratory has explored the utility of supercritical CO_2 extraction as a means of decreasing the time investment for SLB extraction procedures. We have shown that polar phospholipids could not be extracted unless they were prederivatized with trimethylphenylammonium hydroxide (TMPH) to form the methyl esters *in situ* [4]. Super critical fluid extraction, however, did prove remarkably effective in facilitating recovery of DNA from environmental samples[5]. Accelerated solvent extraction (ASE) was used to extract SLB from selected vegetative and/or sporulated biomass as well as from environmental samples collected from water, soil, and air [6]. Dependent on sample type, the automated extraction procedure took approximately 35-45 minutes per sample. Compared to the modified Bligh and Dyer extraction, phospholipid fatty acid (PLFA) lipid yields obtained using the pressurized hot solvent extraction were not significantly different for the vegetative biomass or water and soil samples, but were significantly higher for the spores and the air biomass samples. Initial studies have shown that ASE recovered 3-fold more PLFA from *Bacillus* spores and 2-fold more PLFA from *Aspergillus* spores than the standard one-phase solvent extraction system while significantly reducing extraction time [6]. The system is automated for unattended analysis. The system can be programmed to allow multiple extractions from the same sample so that an automated fractionation can be developed. The increased extraction speed was achieved through a combination of increased pressure and temperature using the Dionex system. The Dionex system can process samples at a temperature of 200°C and pressures of 3,000 psi. with a minimum sample volume of 11 mL (using 18 mL minimum solvent).

Proposed Modifications: The next technological step in rapid extraction will come through the utilization of a system capable of higher pressures such as the ISCO SFX 3560 that will drive the reaction at 10,000psi and 150°C and utilize samples of 0.5 ml volume. This increased pressure will significantly increase the solubility of the lipids & lipophilic membrane proteins from the membranes of the organisms and increase rates of *in situ* derivatization for further separation and analyses. Sufficient separation of "signature" membrane biomarkers will occur rapidly with a high performance liquid chromatography system. Once separated the individual lipid classes and proteins will be ionized by an electrospray ionization source and identified by the ion trap mass spectrometer capable of monitoring successive product ion spectra. The patterns of membrane components could be analyzed by an artificial neural network which will then identify the class or specific type of biological threat

Separation & Detection: The literature is rich in methods for the high performance liquid chromatographic (HPLC) separation and purification of lipids [7]. Moreover, polar lipids readily ionize under

electrospray ionization conditions [8]. To facilitate the initial phases of analysis, a C-18 coated open tubular capillary column between the extraction system and the electrospray ionization ion trap mass spectrometry system will provide in-line separation without the need for post-column splitting. This will provide for the class fractionation and quasi-molecular speciation of the polar lipid components. Electrospray ionization, widely regarded as the most versatile and sensitive method of ionization available, operating in the negative ion mode will provide information regarding the nature of the intact phospholipid, the identity of the two fatty acyl constituents, as well as that of the polar head group. As has been previously demonstrated, organic bases will be added at levels so as to enhance the electrospray signal of phospholipids components [9]. However, pH values greater than 8 have been demonstrated to rapidly deteriorate silica based columns.

The type of mass analyzer which possesses both high sensitivity and a reasonableness as a potential field detector are based on quadrupole based systems. Quadrupole mass spectrometers can be either beam type, with a linear trajectory for the analyte ions, or ion trap quadrupoles, where the ions are stored in a "trap" and ejected for subsequent detection. The ion trap mass spectrometers (ITMS) have features particularly desirable for field instruments such as smaller size and lesser power requirements. The Teledyne ITMS has been chosen by the Army Environmental Center, Aberdeen, MD as its current field portable Tri-service Site Characterization and Analysis Penetrometer System (SCAPS) [10] and seems to be a reasonable choice for the next CBW detector.

Previous investigators have demonstrated the application of fast atom bombardment tandem mass spectrometry to the identification of intact phospholipids. Through a combination of positive ion, negative ion, and constant neutral loss scans, the molecular species of each phospholipid was unambiguously characterized without the need for prior chromatographic purification. Constant neutral loss scans provided assignment of component lipids to their respective phospholipid class (m/z = 141 for phosphatidyl ethanolamine, m/z = 184 for phosphatidylglycerol, etc.). The nature and position of the fatty acyl constituents was determined through daughter scans in the negative ion mode. Preferential formation of the carboxylate anion at the β-position (*sn*-2) results in a daughter ion of much greater intensity than the carboxylate anion at γ-position 1(*sn*-1). Whereas, the positive ion mass spectra provides molecular weight information regarding the identity of the molecular ion minus the polar head group $(M-Y)^+$ [11]. Detection limits specific for phosphatidylethanolamine (PE) were found to be below 1 pg from a mixture of phospholipids. This is roughly the quantity of phospholipids present in 10 *E. coli* cells. In the proposed system, the greater ionization efficiency of electrospray ionization was used to identify phospholipids. Electrospray ionization proved to be an extremely sensitive tool (detection limits below 0.1 pg of material) and a structurally informative means for the rapid identification of intact, underivitized phospholipids. By ramping the cone voltage so as to induce thermal dissociation, fragmentation patterns identical to the tandem collisionally induced techniques were observed in both positive and negative ion mode. The addition of HPLC class and molecular speciation capabilities along with the MS^n features of the ion trap system will result in a rapid, automatable, and field-portable biological warfare detection system [12].

Data Interpretation: Signature lipids offer a great advantage over many other molecular probes in that they reflect the phenotypic expression of the cell's genes. The PLFA in Gram-negative heterotrophs show interesting changes. Cells that age and enter the stationary growth phase show increased ratios of cyclopropane PLFA to their monoenoic PLFA precursors. Cells exposed to metals, solvent, or toxicants form *trans* monoenoic PLFA reversibly from the *cis* monoenoic precursors. This can also result in shifts in the

lipopolysaccharide hydroxy fatty acids. The contractor has utilized this phenotypic expression to show infectiousness of *Cryptosporidium parvum* oocysts [13]. Additionally, proteins offer an enormous potential for identification as they are found in all classes of BW agents. The sequencing of proteins through ES/CID/MS^n allows identification of the proteins specific to potential pathogens.

The most sensitive means to predict outcomes from complex lipid analysis patterns or protein sequences has been to apply artificial neural network to the data [14]. Artificial neural network of patterns of neutral lipid wax components of Mycobacteria were used to "train" the artificial neural network so it could successfully predict from the patterns which were slow growers, which were pathogenic, and which were chromogenic [14]. Artificial neural network is much more powerful than principal components analysis as the artificial neural network can utilize non-linear correlations that are either positive or negative and can utilize additional data as experience to train the system to better predict specific outcomes.

Although signature lipids provide explicit biomarkers for identity and physiological status, most of the signature components encode for contextual information. The same is true for most biological constituents as its function is dependent on the presence of other components [15]. From a data analysis point of view, the entangled interdependencies between analytes requires special tools to recover implicit clinical information. Artificial neural networks provide such a solution by emulating natural learning [16]. Artificial neural network models do not require initial mechanistic assumptions about the biological role of analytes, instead encoding the implicit information from examples ("learning from examples"[17]). Consequently, different artificial neural network can be developed to infer specific traits from a chemical signature.

Data analysis by artificial intelligence techniques will be implemented by first building parallel databases of analytical profiles and microbial identity and characteristics. As database size increases, associations between the two data sets will emerge by continuously integrating the new data in the analysis. The emerging associations will also identify redundant information; therefore help directing data gathering to more interesting samples, in order to attain predictive capability for predefined target parameters. Standard multivariate statistic tools will also be used to describe and test explicit relationships between analytical profiles and corresponding isolates.

Matlab (Mathworks Inc.) provides an adequate software environment to develop artificial neural network applications, being available both for desktop and mainframe operating systems. Applications developed in this environment can be automatically compiled in C, which can then be incorporated directly into the analytical apparatus and provide an immediate evaluation of a new analysis. This solution would provide an intelligent diagnostic tool with the potential to adapt to new situations as they occur. A faulty diagnosis is incorporated as a source of new information ("learning from errors"[17]). Since artificial learning can be continuously implemented, the next diagnosis will have the benefit of recent information.

SLB of threat organisms: The contractor has established that many threat organisms have sufficiently unique signature lipid patterns that are detectable with signature lipid biomarker analysis. Examples include *Bacillus* sp. spores [18], *Francisella tularensis* [19], 17 species of Mycobacteria [14], Legionella [20], and *Cryptosporidium parvum* [13]. Also, lipid patterns can reflect the phenotypic expression of the membrane lipid metabolic capabilities [21]. Information gained can then be used to relate nutritional/physiological status to infectious potential or drug resistance. For example, in *Mycobacterium tuberculosis,* patterns of micoserosic acids and secondary alcohols derived from the surface waxes correlated with drug resistance when subjected to

artificial neural network analysis[13]. With the protozoal pathogen *C. parvum*, signature lipid biomarker analysis using electrospray ionization mass spectrometric allowed detection at the sensitivity of one oocyte, of *C. parvum* based on the unusual PLFA, 10-0H 18:0, found at the β-position of phosphatidylethanolamine [22]. The contractor has established that gas chromatography mass spectrometry systems are also effective, although less sensitive, in the signature lipid biomarker analysis. Making use of current state of the art gas chromatography mass spectrometry instrumentation, this method is capable of detecting as little as 10^3 bacteria [1,13, 21].

CONCLUSIONS: The focus of research could focus on the rapid extraction and detection of both lipids and proteins by developing an ultra-rapid membrane extraction/fractionation system for those microbes and spores accumulated on membrane filters. The filters will be subject to rapid extraction separation lipid and protein profiling and artificial neural network identification of the class or specific biological threat.

References

[1] White, D. C. , D. B. Ringelberg, S. J. Macnaughton, S. Alugupalli, D. P. Schrum. 1996. Signature lipid biomarker analysis for quantitative analysis in situ of environmental microbial Ecology. ACS Symposium 671 "Application of Molecular Markers in Environmental Geochemistry", R. Eagenhouse ed. Pp. 22-34, American Chemical Society, Washington, DC.

[2] Macnaughton S. J., T. L. Jenkins, S. Alugupalli, and D. C. White. 1997. Quantitative sampling of indoor air biomass by signature lipid biomarker analysis: Feasibility studies in a model system. Am Indust Hygiene Assn. J. **58:** 270-277.

[3] Macnaughton, S. J., T. L. Jenkins, M. R. Cormier, R. Gall, and D. C. White. 1997. Lipid biomarker analysis for quantitative analysis of airborne microorganisms. Air & Waste Management Assn., June 8-13 Toronto, CA, CD ROM

[4] Hawthorne, S. B., D. J. Miller, D. E. Nivens, and D. C. White. 1992. Supercritical fluid extraction of polar analytes using *in situ* chemical derivatization. Ann. Chem. **64:** 405-412.

[5] Kehrmeyer, S. R., B. M. Appelgate, H. Pinkert, D. B. Hedrick, D. C. White and G. S. Sayler. 1996. Combined lipid/DNA extraction method for environmental samples, J. Microbiological Methods **25:** 153-163.

[6] Macnaughton, S. J., T. L. Jenkins, M. H. Wimpee, M. R. Cormier, and D. C. White. Rapid extraction of lipid biomarkers from pure culture and environmental samples using pressurized accelerated hot solvent extraction. J. Microbial Methods IN PRESS.

[7] Christie, W. W. 1996. Separation of phospholipid classes by high-performance liquid Chromatography, *In* Advances in Lipid Methodology (W. W. Christie, Ed.), The Oily Press, Dundee, Scotland, Vol 3 pp. 77-107.

[8] Kim, H-Y, T-C. L. Wang, and Y-C. MA. 1994. Liquid chromatography/mass spectrometry of phospholipids using electrospray ionization. Anal. Chem. **66:** 3977-3982.

[9] Wensing, W. W., K. S. Heroux, and A. P. Snyder. 1997. The use of organic bases and acids to enhance the signal of phospholipids by electrospray ionization mass spectrometry. Pp.101. First Tri-Service

Workshop on Biological Mass Spectrometry, Baltimore, July 28-30.

[10] Davis, W. M., J. F. Powell, K. Konecny, J. Furey, C. V. Thompson, M. Wise, and G. Robitaille. 1997. Rapid in-situ determination of volatile organic contaminants in groundwater using the site characterization and analysis penetrometer system. Proceedings of the 1997 U.S.EPA/A&WMA International Symposium on Field Screening Methods for Hazardous and Toxic Chemicals, VIP-XX Air & Waste Management Association, Pittsburgh, PA.

[11] Cole, M. and C. G. Enke. 1991. Direct determination of phospholipid structures in microorganisms by fast atom bombardment triple quadrupole mass spectrometry. Anal. Chem. **63:** 1032-1038

[12] White, D. C., A. A. Arrage, R. S. Burkhalter, C. A. Flemming, K. T. Leung, D. P. Schrum, S. Alugupalli, and S. J. Macnaughton. 1997. A rapid, automatable method to detect and identify the infectious potential of BW agents using signature biomarkers. Pp 23. First Tri-Service Workshop on Biological Mass Spectrometry, Baltimore, MD, July 28-30.

[13] White, D. C., S. Alugupalli, D. P. Schrum, S. T. Kelly, M. K Sikka, R. Fayer and E. S, Kaneshiro. 1997. Sensitive quantitative detection/identification of infectious *Cryptosporidium parvum* by signature lipid biomarker analysis. Proceed. AWWARF conf on Cryptosporidium, Newport Beach CA, March 1-5, 1997.

[14] Almeida, J. S., A. Sonesson, D. B. Ringelberg, and D. C. White. 1995. Application of artificial neural networks (artificial neural network) to the detection of *Mycobacterium tuberculosis*, its antibiotic resistance and prediction of pathogenicity amongst *Mycobacterium spp.* based on signature lipid biomarkers. Binary--Computing in Microbiology **7:** 53-59.

[15] Pahl-Wostl, P. 1995. The dynamic nature of ecosystems. John Wiley & Sons Inc., N.Y., pp. 44-87.

[16] Haykin, S. 1994. Neural networks, a comprehensive foundation. Macmillan Pub., NY, pp22-23.

[17] Hinton, G.E. 1992. How neural networks learn from experience. Sci. American, **267:**144-151

[18] Hedrick, D. B, and D. C. White. 1986. Rapid extraction and fractionation of particulate BW agents for detection of signature biomarkers by mass spectrometry. Proceed. U.S. Army Chem. Defense Res. and Engin. Center Conf. Edgewood Area, APG, MD, Nov. 18-21, p 313-318.

[19] Nichols, P. D., W. R. Mayberry, C. P. Antworth, and D. C. White. 1985. Determination of monounsaturated double bond position and geometry in the cellular fatty acids of the pathogenic bacterium *Francisella tularensis*. J. Clin. Microbiol. **21:** 738-740.

[20] Walker, J. T., A. Sonesson, C. W. Keevil, and D. C. White. 1993. Detection of *Legionella pneumophila* in biofilms containing a complex microbial consortium by gas chromatography-mass spectrometric analysis of genus-specific hydroxy fatty acids.

FEMS Microbiol. Letters **113:** 139-144.

[21] White, D. C. 1995. Chemical ecology: Possible linkage between macro-and microbial ecology. Oikos **74:** 174-181.

[22] Schrum, D.P., S. Alugupalli, S. T. Kelly, D. C. White, R. Fayer, and E. S. Kaneshiro. 1997. Structural Characterization of a "Signature" phosphatidylethanolamine as the Major 10-Hydroxy Stearic acid Containing Lipid of *Cryptosporidium parvum* Oocysts. Lipids **32:** 789-793.

Blank

BIOSYNTHESIS, STRUCTURE AND FUNCTION OF OUTERMEMBRANE LIPOOLIGOSACCHARIDES FROM *NEISSERIA* AND *HAEMOPHILUS*[1]

Bradford W. Gibson[1], Jeffrey Engstrom[1], Arnold Falick[2],
Robert Munson[3], Anthony Campagnari[4], and Michael Apicella[5]

[1]University of California, San Francisco
School of Pharmacy
San Francisco, California

[2]PerSeptive Biosystems
South San Francisco, California

[3]Children's Hospital Research Foundation
Columbus, Ohio

[4]SUNY
Buffalo, New York

[5]University of Iowa
Iowa City, IA

ABSTRACT

The outer-membrane of pathogenic gram-negative *Neisseria* and *Haemophilus* species contain proteins, polysaccharides and glycolipids that play critical roles in the organism's ability to colonize human tissue and evade host defenses. To understand the role of the surface gylcolipids or lipooligosaccharides (LOS) in the disease process, we have been developing methods for their structural and functional characterization. LOS are highly complex and consist of hydrophobic and anionic Lipid A moiety, which serves to embed the LOS in the outer-membrane, and a complex oligosaccharide that is linked to the lipid A moiety through the acidic sugar 2-keto-3 deoxyoctulosonic acid. LOS contain both neutral and acid sugars as well as variable amounts of phosphate and phosphoethanolamine substituents. Our laboratory has developed both electrospray and matrix-assisted laser desorption ionization (MALDI) techniques for the analysis of LOS. To date, we have used these methods to determine the LOS structures from a variety of *Haemophilus* strains and have begun to piece together the biosynthetic repertoire of these bacteria. In this report, we will describe how mass spectrometry has been used to provide details of LOS structure and how it has provided crucial information as to the biological roles of these glycoconjugates in the disease process.

[1] Support for this work was provided by grants from the NIH (AI 31254 and AI 24616) and National Center for Research Resources (RR 01614).

Blank

ANALYSIS OF GENETIC POLYMORPHISMS IN PCR AMPLIFIED DNA BY MALDI-TOF MASS SPECTROMETRY

Philip Ross and Phillip Belgrader
DNA Technology Branch, Center for Medical and Molecular Genetics
Armed Forces Institute of Pathology
Rockville, Maryland

ABSTRACT

Matrix assisted laser desorption-ionization time of flight mass spectrometry (MALDI-TOFMS) is an emerging technology for size and sequence analysis of amplified DNA fragments. The presence of reaction components encountered in DNA amplification prohibits effective MALDI-TOFMS analysis. We have developed rapid and efficient methods for DNA analysis which circumvent the rigorous time consuming purification procedures typically required for successful mass spectrometric analysis. These novel approaches are demonstrated for amplified DNA containing repeat polymorphisms and point mutations. In one approach, PCR amplified DNA fragments are captured, washed and directly analyzed by MALDI-TOFMS, yielding accurate molecular weight information for fragments over 100 bases in length. Alternatively, an assay using peptide nucleic acid (PNA) probes recognizing specific regions harboring single base mutations has been developed. The PNA probes are detected with high mass accuracy by MALDI-TOFMS following rapid, sequence specific hybridization to immobilized DNA. Both methodologies are performed in a single reaction tube using simple operations and therefore are directly amenable to automated processing.

Blank

PROTEIN FOLDING AND INTERACTIONS STUDIED BY MASS SPECTROMETRY

Carol V. Robinson
Oxford Centre for Molecular Sciences,
New Chemistry Laboratory,
South Parks Road,
Oxford OX1 3QT, UK.

ABSTRACT

The conformational properties of proteins have been measured by a combination of hydrogen exchange and electrospray mass spectrometry. These studies provide insight into the stability and conformational dynamics of proteins as well as allowing investigation of the folding pathway of proteins. The effects of molecular chaperones on folding have also been investigated *via* gas phase dissociation of large multi-protein complexes and measurement of hydrogen exchange properties of the protein ligand. These techniques have also been applied to the study of the interactions and conformations of proteins which misfold *in vivo* and give rise to amyloid fibrils.

1. INTRODUCTION

The role of mass spectrometry (MS) in elucidating the primary structure of proteins is well established. The advent of electrospray interfaces capable of analyzing proteins from solution conditions under which the protein remains in its native state have opened the door to MS studies which reflect the solution state of proteins. These MS experiments rely on retention of the 3D structure of proteins until the protein is in the gas phase of the mass spectrometer.

Previously we have reported mass spectrometry studies which have allowed the detection of a transient folding intermediate in the folding of avian and human lysozymes [1] [2], the cooperative folding of a 3 disulfide hen lysozyme derivative (CM^{6-127}) [3] and an acyl CoA binding protein (ACBP) [4]. An extension of these studies to include the role of molecular chaperones have revealed the conformation of the protein ligand bound on the surface of the GroEL oligomer [5]. In this article I will review the role of mass spectrometry for studying protein conformation in protein - ligand and protein - protein complexes and the conformational dynamics of two variants of human lysozyme involved in amyloid fibril formation.

2. EXPERIMENTAL

All mass spectra were collected on Micromass Platform mass spectrometers from 100% aqueous conditions in the absence of organic co-solvent and without source heating. Hydrogen exchange was initiated by dilution of protein containing solution in H_2O into D_2O at the appropriate pH. Formic acid was used to adjust the pH. All pH measurements are glass electrode meter readings and no attempt has been made to correct for solvent isotope effects. Protein samples were prepared as described elsewhere [6].

3. RESULTS AND DISCUSSION

3.1 HYDROGEN - DEUTERIUM EXCHANGE

The hydrogen exchange kinetic profiles of native hen lysozyme and a 3 disulfide derivative of hen lysozyme in which a critical disulfide bond has been reduced and carboxymethylated (CM^{6-127}) are shown in figure 1. Both proteins where diluted 100 fold into D_2O at pH*3.8 and their change in average mass monitored as a function of time. The faster exchange rate observed in the 3-disulfide derivative compared

with the native protein reflects the reduced overall stability of the protein under these conditions [3]. The same protein samples were investigated by 1D-^{1}H NMR under the same solution conditions. Despite the different concentrations required for the two different techniques the results were found to be identical within the errors of the techniques[6].

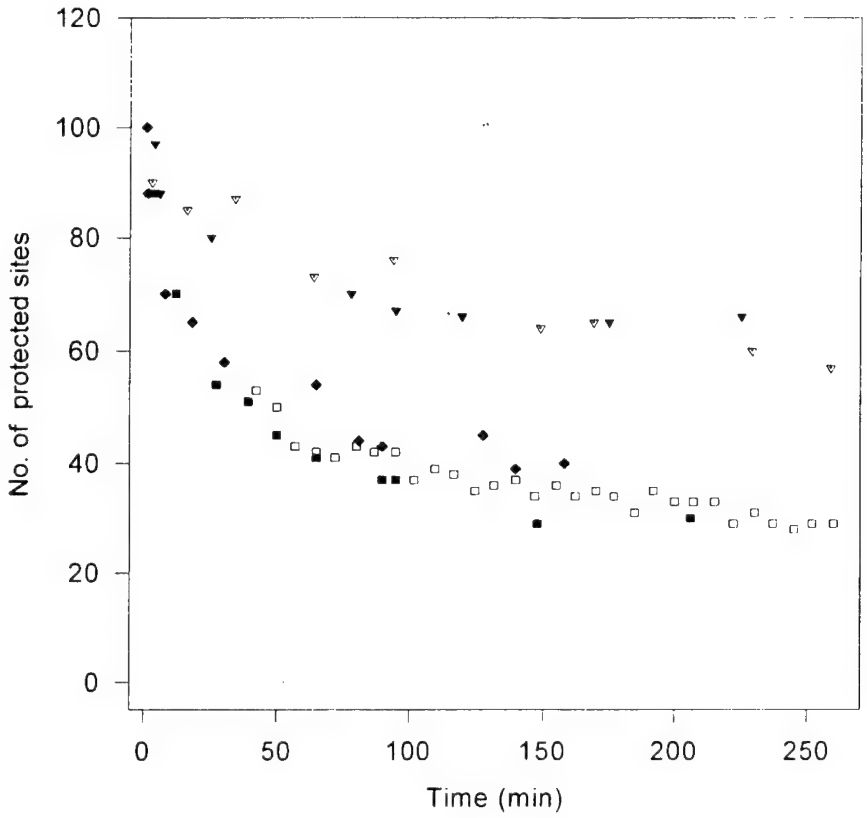


Figure 1. Comparison of the hydrogen exchange kinetic profiles of native hen lysozyme (▼) at pH5 by mass spectrometry (filled symbols) and NMR (open symbols) with $CM^{6\text{-}127}$ lysozyme (◆) measured by mass spectrometry demonstrating the reduced protection against exchange in this 3 disulphide derivative. A comparison between the mass spectrometry data and NMR data (■ and □ respectively) of $CM^{6\text{-}127}$ lysozyme was also carried out at pH3.8 demonstrating the consistency of these measurements with corresponding NMR solution experiments.

3.2 PROTEIN - LIGAND COMPLEXES

The mass spectrum of an acyl CoA binding protein (ACBP) from standard electrospray conditions (50% acetonitrile, 1% formic acid and source temperature of 50°C) is shown in figure 2 (upper trace). These conditions are used routinely in ESI MS to obtain high quality spectra of proteins to confirm their molecular weight.

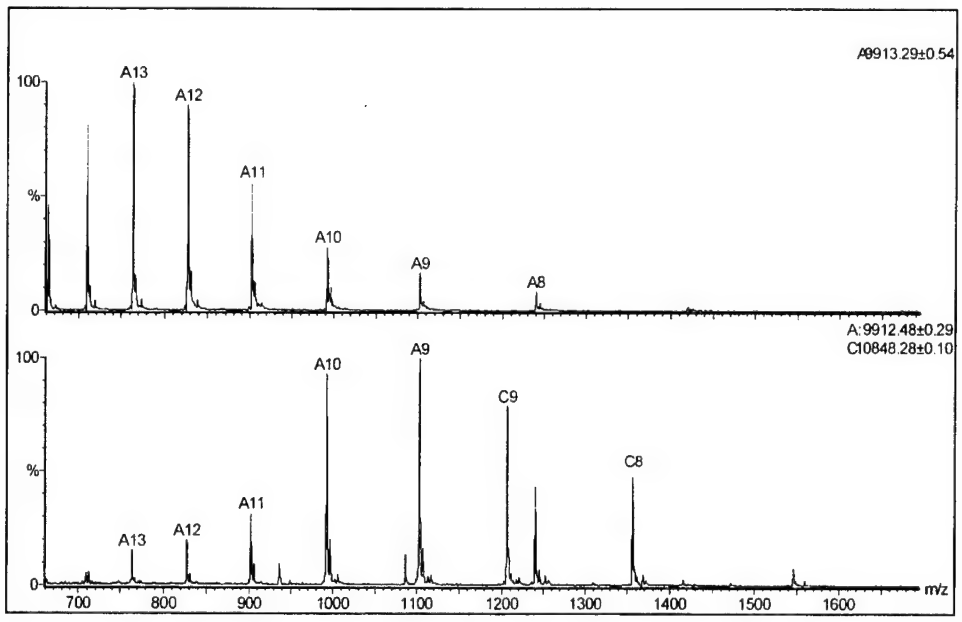


Figure 2 . Electrospray mass spectra of ACBP from 50% acetonitrile and 1% formic acid (upper trace) and 100% water at pH5 (lower trace). The change in charge state distribution, from a maximum of +13 to +9 and the additional charge state series labelled C, corresponding in mass to non-covalent complex formed between ACBP and a CoA ligand, reflect the fact that the protein was analysed from solution in which it remained tightly folded.

When the solution conditions are adjusted to 100% water at pH5 and the spectrum obtained, the protein is in its native folded conformation, figure 2 lower trace. Comparison of the two spectra from different solution conditions reveal a shift in the charge states and the appearance of a new series of peaks with a higher mass corresponding to protein - CoA ligand complex. This shift in charge state and appearance of the non-covalent complex arise since the proteins is in its native folded conformation in solution.

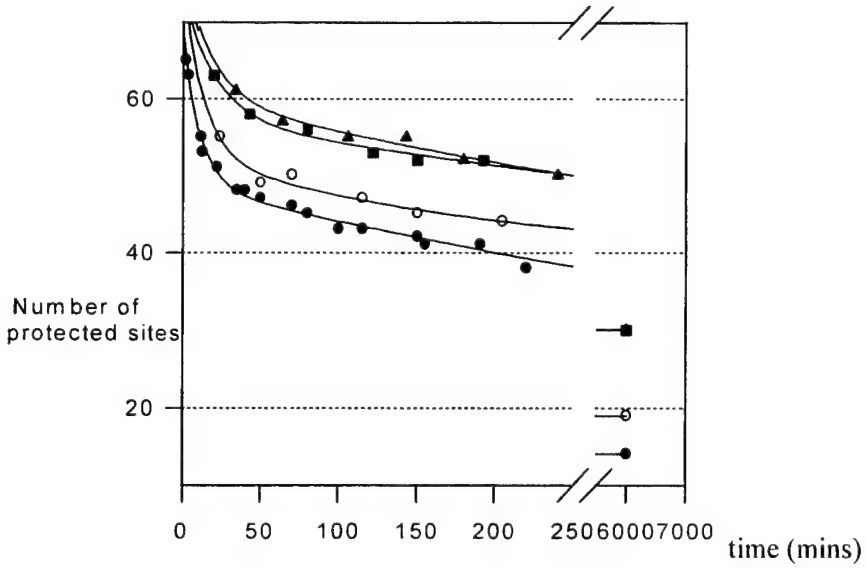


Figure 3. Hydrogen exchange kinetic profiles of ACBP at pH5 and 20°C in the presence of tightly binding octanoyl (■) and dodecanoyl CoA ligand (▲), weaker binding acetyl CoA ligand (o) and in the absence of ligand (●). The increased hydrogen exchange protection afforded by the tighter binding CoA ligands compared with the weaker binding ligand and the *apo* protein is clearly demonstrated.

It is interesting to note that whether the protein-ligand complex remained intact or dissociated in the gas phase the same number of deuterons remained with the protein. Thus the protein carries a signature of its solution conformation by virtue of its hydrogen exchange properties. This ability to dissociate large protein complexes while retaining their hydrogen exchange history can be used to investigate much larger protein assemblies, see below.

3.3 CHAPERONE ASSISTED PROTEIN FOLDING

The folding of nascent polypeptide chains *in vivo* is mediated through interactions with molecular chaperones. We have developed a method for monitoring the conformation of the protein ligand in the presence of the chaperonin GroEL. The method is based upon dissociation of the complex and measurement of the hydrogen exchange properties of the ligand on the surface of the GroEL complex, figure 4, [8]. GroEL is a multimeric protein complex consisting of 14 identical 57kDa subunits. Gas phase dissociation and measurement of the deuterium content of the protein ligand enables different folding intermediates bound to GroEL to be examined. In a recent study of human dihydrofolate reductase bound to GroEL two distinct folding intermediates exhibited closely similar hydrogen exchange properties. These results lead to the suggestion that chaperone mediated folding occurs by minor modification of the protein ligand on the surface of the GroEL complex rather than by progressive folding to the native state or complete unfolding on the surface of GroEL [9]

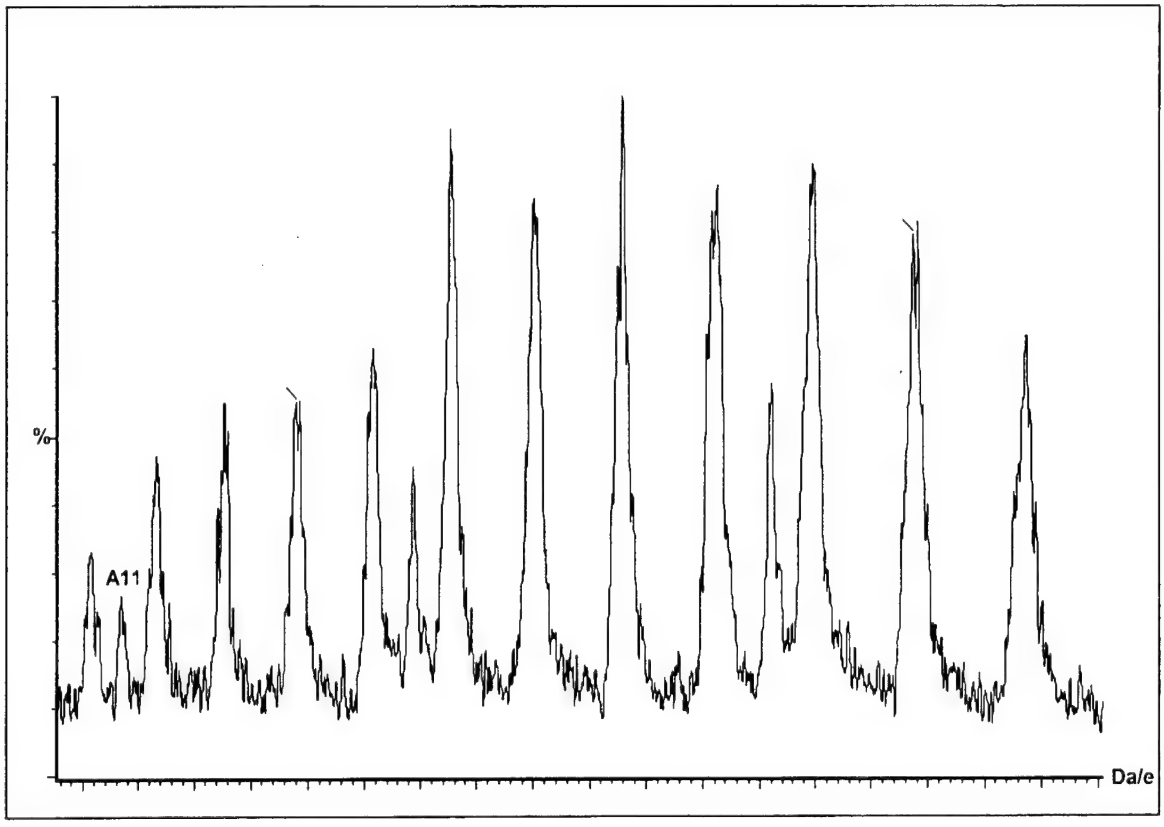


Figure 4. Electrospray mass spectrum of the complex formed between GroEL and bovine α-lactalbumin showing dissociation in the gas phase to give two charge state series, series A arising from the α-lactalbumin ligand and series B from GroEL monomers. This gas phase dissociation allows the hydrogen exchange properties of the protein ligand to be measured without dissociation in solution prior to analysis.

3.4 PROTEIN MISFOLDING

Protein misfolding underlies a range of fatal disorders including Alzheimer's disease and transmissible spongiform encephalopaphy. The proteins involved in these diseases have unrelated sequences yet can all polymerize into fibrils with similar ultrastructural properties. Using mass spectrometry we have investigated the hydrogen exchange properties of two variants of human lysozyme the Asp67His and Ile56Thr and compared these properties with wild type human lysozyme under identical conditions, [10] figure 5. The two amyloid forming variants of lysozyme display a dramatic decrease in hydrogen exchange protection when compared to wild type human lysozyme. These results are interpreted in terms of increased conformational dynamics of the protein variants allowing access to the hydrophobic core of the protein thus facilitating hydrogen exchange. These studies, together with other biophysical data, allowed a proposed mechanism in which the mainly helical fold of the protein, observed in crystal structures, is in equilibrium with a partially folded form capable of conversion to amyloid fibril.

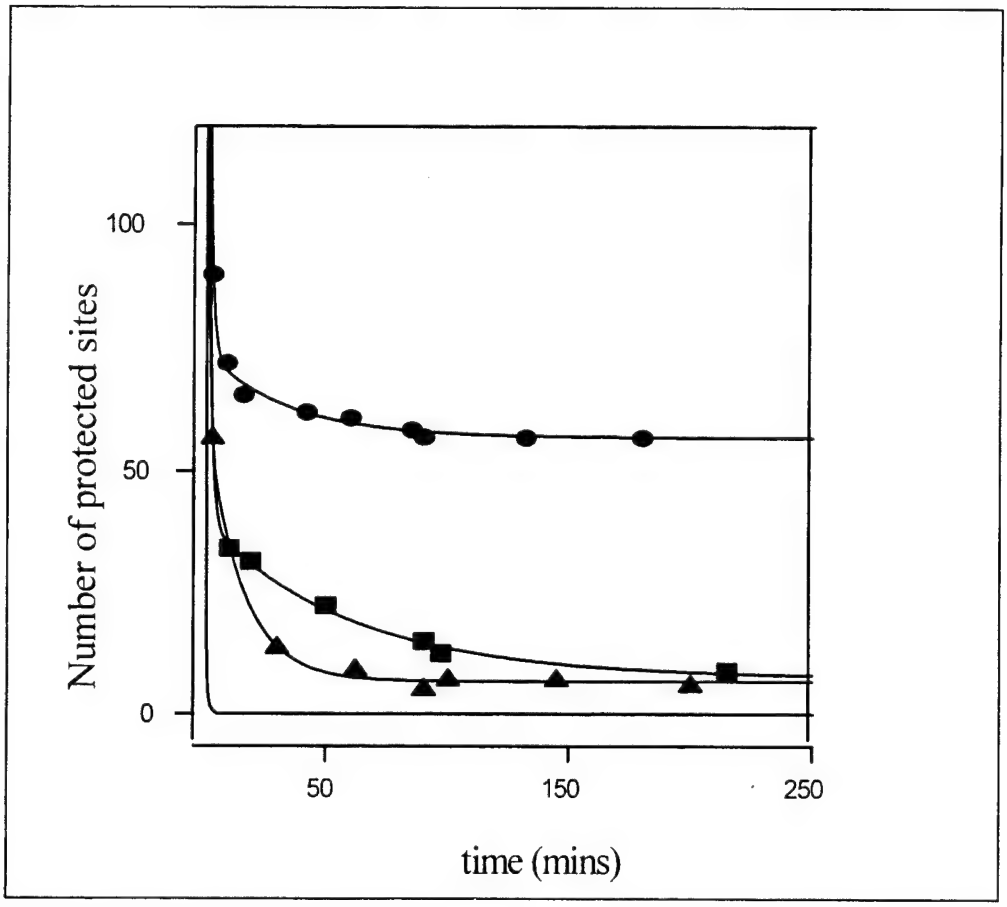


Figure 5. Hydrogen exchange kinetic profiles of wild type human lysozyme (•), Asp67His (■) and Ile56Thr (▲) at pH5 and 37°C. The dramatic difference in exchange rates observed in the variants is attributed to increased conformational dynamics occurring in solution which allow access to the hydrophobic core of the protein and thus facilitate hydrogen exchange.

4. CONCLUSION

The utility of the hydrogen exchange process resulting from its sensitivity to changes in solution conditions and protein conformation is exemplified in this study. Comparison of the MS data with NMR data under the same solution conditions show agreement within the experimental errors of the two techniques. The different hydrogen exchange rates observed for protein in the presence of different ligands and the

conformation of protein ligands bound in multimeric protein complexes have provided important insight into these biological systems. Moreover, the ability to examine increased conformational dynamics of amyloidogenic proteins has provided key insight into the mechanism of amyloid fibril formation *in vivo*.

5. REFERENCES

1 Miranker, A., Robinson, C. V., Radford, S. E., Aplin, R. T. and Dobson, C. M. (1993) Science 262, 896-899.

2 Hooke, S. D., Eyles, S. J., Miranker, A., Radford, S. E., Robinson, C. V. and Dobson, C. M. (1995) J. Am. Chem. Soc. 117, 7548-7549.

3 Eyles, S. J., Radford, S. E., Robinson, C. V. and Dobson, C. M. (1994) Biochemistry 33, 13038-13048

4 Kragelund, B. B., Robinson, C. V., Knudsen, J., Dobson, C. M. and Poulsen, F. M. (1995) Biochemistry 34, 7217-7224.

5 Robinson, C. V., Groβ, M. and Radford, S. E., (1997) Methods in Enzymology in press

6 Chung, E. W., Nettleton, E. J., Morgan, C. J., Grob, M., Miranker, A., Radford, S. E., Dobson, C. M. and Robinson, C. V. (1997) Protein Science 6, 1316-1324.

7 Robinson, C. V., Chung, E. W., Kragelund, B. B., Knudsen, J., Aplin, R. T., Poulsen, F. M. and Dobson, C. M. (1996) J. Am. Chem. Soc. 118, 8646-8653.

8 Robinson, C. V., Groβ, M., Eyles, S. J., Ewbank, J. J., Mayhew, M., Hartl, F. U., Dobson, C. M. and Radford, S. E. (1994) Nature 372, 646-651.

9 Groβ, M., Robinson, C. V., Mayhew, M., Hartl, F. U. and Radford, S. E. (1996) Protein Science 5, 2506-2513.

10 Booth, D., Sunde, M., Bellotti, V., Robinson, C. V., Hutchinson, W. L., Fraser, P. E., Hawkins, P. N., Dobson, C. M., Radford, S. E., Blake, C. C. F. and Pepys, M. B. (1997) Nature 385, 787-793.

IDENTIFYING PROTEIN-NUCLEIC ACID CONTACTS BY PHOTOCHEMICAL CROSSLINKING AND MASS SPECTROMETRY

Douglas F. Barofsky
Department of Agricultural Chemistry and
Environmental Health Science Center
Oregon State University
Corvallis, Oregon 97331, USA

Elucidation of the molecular interactions between proteins and nucleic acids is considered essential to understanding the basic cellular processes involved in the replication, recombination, dynamic alteration, and repair of DNA and in the transcription and regulation of genes. In order to fully understand the interactions between proteins and nucleic acids at the molecular level, it is necessary to identify the amino acid and nucleotide residues in addition to determining the nature and magnitude of the driving forces specifically involved in protein-nucleic acid interactions. Recently, Jensen et al.[1, 2] described a protocol for using matrix-assisted laser desorption/ionization (MALDI) mass spectrometry[3] and electrospray ionization (ESI) mass spectrometry[4] to investigate nucleic acid binding proteins photochemically crosslinked to nucleic acid monomers or oligomers.

It has been known for some time that the interaction between a protein and a nucleotide can be fixed by using ultraviolet (UV) light to induce a "zero-length," covalent bond between the two biopolymers.[5] Although the mechanisms governing the formation of the covalent bonds are not fully understood, techniques based on this phenomenon have proven useful in exploring both static and dynamic protein-nucleic acid complexes.[6, 7] When performed under proper experimental conditions, UV-induced crosslinking apparently only takes place with native protein and only involves the amino acid residues that are proximal to photo-excited pyrimidines in the nucleic acid molecule. Thus, identification of the modified amino acids and the nucleobases they are crosslinked to can provide information about the binding domains of the protein and nucleic acid respectively. Identifications of this sort can be accomplished by mass spectrometric methods.

The strategy outlined by Jensen et al. can be broken down into three stages.[2] In the first stage, the nucleic acid binding protein and its nucleic acid substrate, or a photoactivatable analogue thereof, are incubated under proper conditions to form the protein-nucleic acid complex. The sample is then irradiated with UV-light for a given period of time in order to create photochemically crosslinked protein-nucleic acid complexes. The irradiated reaction mixture is analyzed either in an unpurified or purified state by MALDI mass spectrometry to establish the molecular weight of the crosslinked product. If the crosslinking yield is low, the crosslinked protein-nucleic acid complex can be separated from the unreacted reagents before further characterization.

In the second stage, the crosslinked complex is digested with a protease, a nuclease, or both, and the resulting mixture is mass spectrometrically analyzed in order to generate mass specific peptide or oligonucleotide maps. Comparison of mass spectrometric maps of the protein or nucleic acid produced before and after UV-crosslinking can be used to identify peptides or oligonucleotides that deviate from their expected molecular weights by a mass increase corresponding to a nucleic acid or peptide addition respectively. This strategy relies, of course, on the fact that, in most cases, the amino acid sequence of the protein and nucleotide sequence of the nucleic acid is known, and it will, therefore, be possible to predict the respective molecular weights of the peptides and oligonucleotides resulting from cleavages by a specific enzyme.

In the third stage, the modified peptides or oligonucleotides identified in the second stage are sequenced by tandem mass spectrometry in order to identify the amino acid or nucleotide residues involved in forming the photochemical crosslink and, thereby, the binding sites of the protein or nucleic acid respectively.

To date, there have been at least four applications of this protocol reported in the literature. Each is summarized below.

URACIL-DNA GLYCOSYLASE

Eserichia coli uracil DNA glycosylase (Ung) is a monomeric protein consisting of 228 amino acids whose enzymatic role is to initiate the uracil-excision DNA-repair pathway by removing a uracil-base introduced into

DNA.[8] MALDI mass analysis was used extensively in a study of the DNA binding site of Ung to identify nucleopeptides in an enzymatic digest of a UV crosslinked Ung-dT_{20} complex.[9] Results from the MALDI mass analyses were combined in this study with concurrent results from Edman sequencing and biological testing to demonstrate that amino acids 58-80 (tryptic peptide 6) and amino acids 185-213 (tryptic peptides 18 and 19) reside at the interface of an Ung-DNA complex. These two domains in the primary sequence of Ung, which are heavily conserved in general, are especially similar to the corresponding amino acid regions in human uracil-DNA glycosylase (UNG 15). Viewing their findings for Ung in light of this fact, Bennett et al. ventured that the two UNG 15 amino acid residues Asp145 (corresponding to Asp64 in Ung) and His268 (corresponding to His187 in Ung), whose substitution was recently shown by site-directed mutagenesis studies on the human uracil-DNA glycosylase gene to inactivate UNG 15,[10] reside in the vicinity of the DNA-binding pocket.[9] This speculation has since been substantiated by the 3-dimensional structure of UNG 15 obtained from x-ray analysis.[11] Another recent x-ray diffraction structure, this one of herpes simplex virus uracil-DNA glycosylase,[12] further corroborates the general validity of the Ung-findings by Bennett et al. and the utility of the mass spectrometric approach they used to obtain them.

SYNTHETIC NUCLEOPEPTIDES

Peptide-oligonucleotide heteroconjugates are synthesized for use in antisense research; the peptidic components of these hybrid complexes are thought to help stabilize, transport, or target antisense oligonucleotides.[13] In order to further develop mass analysis as a tool for investigating covalent nucleopeptides, such as those encountered in antisense research and in studies that use photochemical crosslinking to investigate molecular aspects of protein-nucleic acid interactions, Jensen et al. have chemically synthesized peptide-oligothymidylic acid conjugates and examined their behavior under MALDI and ESI mass spectrometric conditions.[14] Two peptide-oligothymidylic acids, prepared by joining an eleven-residue synthetic peptide containing one internal carboxyl group (Asp side chain) to amino-linker-5′pdT_6 and amino-linker-5′pdT_{10} oligonucleotides, were analyzed by MALDI on a linear time-of-flight mass spectrometer and by ESI on a triple-quadrupole system. MALDI and ESI sensitivities for the two hybrid compounds were found to be similar respectively to their sensitivities for the pure oligonucleotide parts. In general, MALDI proved to be less affected by sample impurities and more sensitive than ESI, while ESI on the quadrupole produced greater mass accuracy and resolution than MALDI on the time-of-flight instrument. A hybrid's behavior in a MALDI-matrix or an ESI-spray-solvent was found to be governed mainly by the oligonucleotide. A single positive ESI tandem mass spectrum of the peptide-dT_6 accounted for the heteroconjugate's entire primary structure including the point of the oligonucleotide's covalent attachment to the peptide.

The tendencies observed with the two synthetic peptide-oligonucleotide conjugates examined in this investigation provide a starting point for future, systematic studies of the relative contributions of oligopeptides and oligonucleotides to the physicochemical properties of heteroconjugates. From a practical perspective, the study points to new ways to productively employ mass spectrometry in antisense research and in studies of protein-nucleic acid interactions, namely, monitoring the synthesis or UV-crosslinking of oligopeptide-oligonucleotide hybrids, confirming molecular masses of hybrids, and obtaining residue-specific information pertaining to sites of conjugation.

REC A

Rec A is a multifunctional protein with an M_r of about 38,000. In recombination it acts to promote pairing of homologous strands. In a study of the protein binding sites of E. coli RecA to DNA, Qin and Chait photocrosslinked a 50mer stretch of DNA (in which some of the thymine bases were substituted by photoreactive iodouracil) with RecA protein by irradiating with the 308 nm output from a XeCl excimer laser.[15] The resulting reaction mixture was digested with trypsin, and the photocrosslinked peptide-DNA fragment purified by anion ion exchange chromatography. The photocrosslinked peptide was then digested with Nuclease P1 (a nonspecific nuclease that digests DNA completely leaving a single modified nucleotide bound to the peptide of interest), and the product was analyzed by MALDI-ion trap MS. No attempt was made to identify the crosslinked tryptic peptide by measuring the mass difference between the crosslinked peptide and the unmodified peptide because neither the

photochemistry nor the gas phase ion dissociation chemistry of photocrosslinks in an ion trap mass spectrometer have been fully described. Instead, MS/MS was used to make the identification. The MS/MS spectrum of an ion with m/z 1861.8 exhibited four strong ion fragments in addition to small neutral losses from this precursor. Taking advantage of the facts that 1) the products were the result of preferential cleavage at the C-termini of acidic residues, 2) the known sequence of RecA, and 3) trypsin was used for the proteolytic digestion, it was deduced that the crosslinked peptide was 154AEIEGEIGDSHMGLAAR170. MS/MS/MS carried out on the MS/MS fragment at m/z 947.8 established that the crosslinked amino acid residue was 165Met. Once the crosslinked peptide and the mass of the remaining crosslinker was established, it was possible to ascertain the identity of the crosslinked peptide, the crosslinked amino acid residue, and the likely site of crosslinking on the sulfur atom of methionine. These mass spectrometric assignments were independently confirmed through Edman sequencing of the crosslinked peptide where release of the modified methionine yielded a silent sequencing step.

RIBOSOMAL PROTEINS

A considerable amount of structural data for proteins complexed with DNA has been amassed. By comparison, only a few structures of RNA-protein complexes have been solved. Urlaub et al. have successfully applied a version of Jensen et al.'s protocol to an investigation of RNA-protein interaction.[16] In this study, crosslinked peptide-oligoribonucleotide complexes derived from distinct regions of the rRNA and the individual ribosomal proteins of the 30 S ribosomal subunits from *Escherichia coli* were isolated and purified. N-terminal amino acid sequencing in conjunction with MALDI time-of-flight mass spectrometry was employed to determine the crosslinking sites at both the amino acid and nucleotide levels. The MALDI analyses were performed following partial alkaline hydrolysis of the crosslinked peptide-oligoribonucleotide complexes. It was found that Lys-44 in S4 was crosslinked to the oligoribonucleotide to either U-1541 or A-1542 in the 16 S RNA, Lys-75 in S7 to C-1378, Met-114 in S7 to U-1240, Lys-55 in S8 to U-653, and Lys-29 in S17 to U-632. The fact that the nucleotide sites belong to non-base paired regions of the 16 S RNA, adds direct structural evidence that loop structures within ribosomal proteins together with looped or bulged structures within the rRNA are crucial for protein-RNA interactions.

The preceding examples demonstrate that the speed, sensitivity, and specificity of MALDI mass spectrometry and ESI mass spectrometry used in conjunction with photochemical crosslinking greatly increase the range of experiments that are possible for investigating structural aspects of protein-nucleic acid interactions at the molecular level.

REFERENCES

[1]O. N. Jensen, D. F. Barofsky, M. C. Young, P. H. von Hippel, S. Swenson, and S. E. Seifried, Rapid Commun. Mass Spectrom. **7**, 496 (1993).

[2]O. N. Jensen, D. F. Barofsky, M. C. Young, P. H. von Hippel, S. Swenson, and S. E. Seifried, in *Techniques in Protein Chemistry V*, edited by J. W. Crabb (Academic Press, San Diego, 1994), p. 27.

[3]M. Karas, D. Bachmann, U. Bahr, and F. Hillenkamp, Int. J. Mass Spectrom. Ion Processes **78**, 53 (1987).

[4]C. K. Meng, M. Mann, and J. B. Fenn, Zeitschrift für Physik D – Atoms, Molecules and Clusters **10**, 361 (1988).

[5]M. D. Shetlar, in *Photochemical and Photobiological Reviews*, edited by K.C. Smith (Plenum Press, New York, 1980), Vol. 5, p. 105.

[6]K. R. Williams and W. H. Konigsberg, in *Protein-DNA Interactions*, edited by R. T. Saurer (Academic Press, San Diego, 1991), Vol. 208, p. 516.

[7]J. W. Hockensmith, W. L. Kubasek, W. R. Vorachek, E. M. Evertsz, and P. H. von Hippel, in *Protein-DNA Interactions*, edited by R. T. Saurer (Academic Press, San Diego, 1991), Vol. 208, p. 211.

[8]D. W. Mosbaugh and S. W. Bennett, Prog. Nucleic Acid Res. Mol. Biol. **48**, 315 (1994).

[9]S. F. Bennett, O. N. Jensen, D. F. Barofsky, and D. W. Mosbaugh, J. Biol. Chem. **269**, 21870 (1994).

[10]H. E. Krokan, G. Slupphaug, I. Alseth, B. Kavli, T. Haug, L. C. Olsen, and D. E. Helland, in *New York Academy of Sciences*, Burlington, VT, (1993), abstract.

[11]C. D. Mol, A. S. Arvai, G. Slupphaug, B. Kavli, I. Alseth, and H. Krokan, Cell **80**, 869 (1995).

[12]R. Savva, K. McAuley-Hecht, T. Brown, and L. Pearl, Nature **373**, 487 (1995).
[13]J. S. Cohen and M. E. Hogan, Sci. Am. **271**, 50 (1994).
[14]O. N. Jensen, S. Kulkarni, J. V. Aldrich, and D. F. Barofsky, Nucleic Acids Res. **24**, 3866 (1996).
[15]J. Qin and B. T. Chait, Anal. Chem. (in press).
[16]H. Urlaub, B. Thiede, E.-C. Müller, R. Brimacombe, and B. Wittmann-Liebold, J. Biol. Chem. (in press).

FAST IDENTIFICATION OF BIOLOGICAL AGENTS USING NANOFLOW ELECTROSPRAY TIME OF FLIGHT (TANDEM) MASS SPECTROMETRY

Peter VERHAERT, Arnold DE LOOF
Comparative Animal Physiology and Morphology Group,
University of Leuven, Department of Zoology,
Naamsestraat 59, B-3000 LEUVEN (Belgium)

Mike MORRIS & Bob BORDOLI
Micromass UK Ltd, Floats Road,
Wythenshawe, Manchester M23 9LZ (UK)

This paper addresses some features of hybrid quadrupole orthogonal time of flight (TOF) technology in combination with nano-electrospray ionization (ESI), which in our opinion render this technology exceptionally powerful with respect to the central topic of this First Joint Services Workshop on Biological Mass Spectrometry. Particularly in combination with a proper "biologically driven" selection of the analyte, sample preparation efforts and -time may be reduced to an absolute minimum.

The potential of the technique is illustrated with an example from comparative neurobiology, where this highly sensitive and rapid biological MS and MS/MS method allows one to unambiguously identify species-specific neurosecretory products, within minutes, from as little sample as a tiny part of neuroendocrine tissue from an insect nervous system.

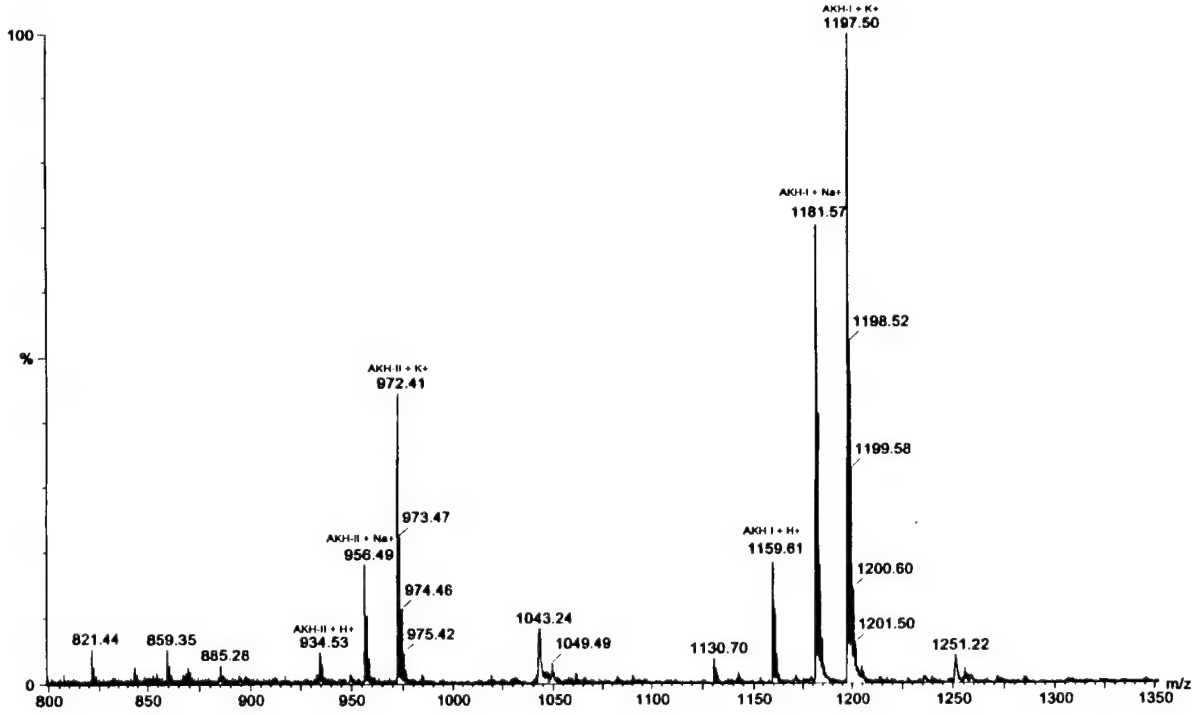
At this workshop particular emphasis has been put on the recent technological advances of mass spectrometry (MS) to detect and identify various agents of biological origin (ABO: bacteria, viruses or proteinaceous toxins). Especially developments towards an improvement of its ease of use, speed and efficacy (sensitivity and reliability) have been of particular interest, as these features are considered essential if the technique is ever to be implemented in a battlefield context [1].

Conventional techniques requiring elaborate extraction procedures are not entirely suitable in view of the main objective of this Workshop [2]. It is, therefore, unfortunate that most, if not all, of the biomarker molecules which one is currently focusing upon to identify the different types of biological warfare (BW) agents require more or less extensive experimental sample manipulations, prior to their introduction into the MS analysis system [see 3 for an overview]. However some recent data obtained by various groups indicate that "species-specific" MS profiles can be generated from entire, undisrupted and unextracted viruses [4] as well as bacteria [5], respectively by use of ESI triple quadrupole and matrix-assisted laser desorption/ionization (MALDI) TOF equipment.

Having recently introduced the newly developed hybrid ESI-TOF technology in neurobiology, we feel that this type of instrumentation will even further improve the above MS identification methods [6]. With this technique it is made possible to obtain species-specific MS profiles of the molecules secreted by small pieces of living tissue, and at the same time acquire MS/MS information of individual peaks contained therein. In our model system, the insect corpus cardiacum (a neuronal

source of a variety of neuropeptides), it was sufficient to bathe the miniscule organ from a single animal for as little as 2-5 minutes in 10 µl of a (1:1) methanol:0.5% formic acid solution, before introducing it into the MS source. For this, 1 µl of the sample was deposited (using an Eppendorf GeLoader) directly into the gold coated borosilicate nanoflow ESI capillary which was then fitted to the sample probe as described below. Acquisition of a full scan MS profile of the tissue followed within a few minutes and the 1 µl amount of sample allowed further complete MS/MS analysis of, at least, two different peaks from the MS spectrum (unequivocally identifying the amino acid sequence of the respective neuropeptides). The entire process, including tissue disection, took less than 45 minutes.

Two insect species were compared in this study, the American cockroach, *Periplaneta americana* and the African desert locust, *Schistocerca gregaria*. The 900-1200 Da mass range appeared particularly interesting. The two heptapeptide cardioacceleratory hormones of the cockroach (PeaCAH I & II [7;8]) and the two adipokinetic hormones of the locust (Scg AKH I & II, a deca- [9] and an octapeptide [10]) clearly showed up as (M+H)+, (M+Na)+ and (M+K)+ ions (Fig. 1).



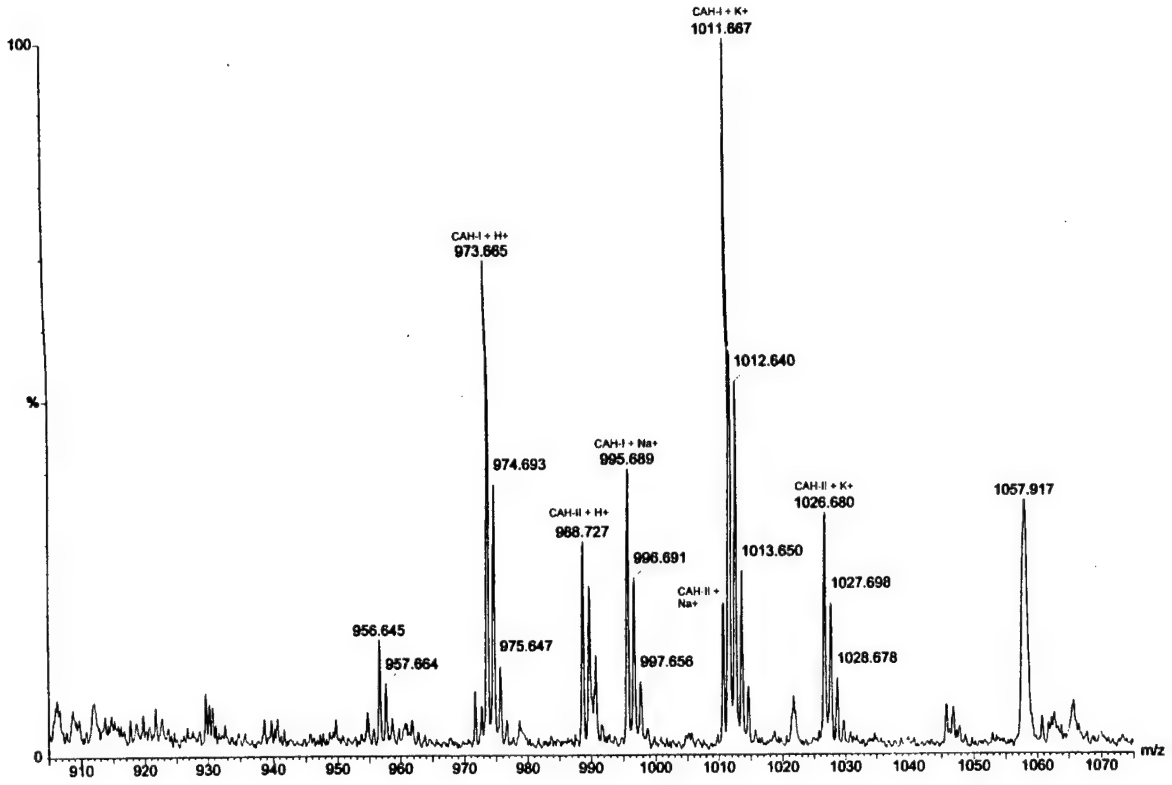


Figure 1. Q-Tof MS profiling of neuropeptides from the corpus cardiacum of a locust (a) and a cockroach (b). Between 900 and 1200 dalton different sets of MS peaks are obvious, allowing the insect species to be identified. CAH I, pGluValAsnPheSerProAsnTrpamide; CAH II, pGluLeuThrPheThrProAsnTrpamide; AKH I, pGluLeuAsnPheThrProAsnTrpGlyThramide; AKH II, pGluLeuAsnPheSerThrGlyTrpamide.

Tandem MS analysis on the protonated ion species of all four peptides yielded unequivocal sequence data (both y"- and complementary b-ions; see example in Fig. 2).

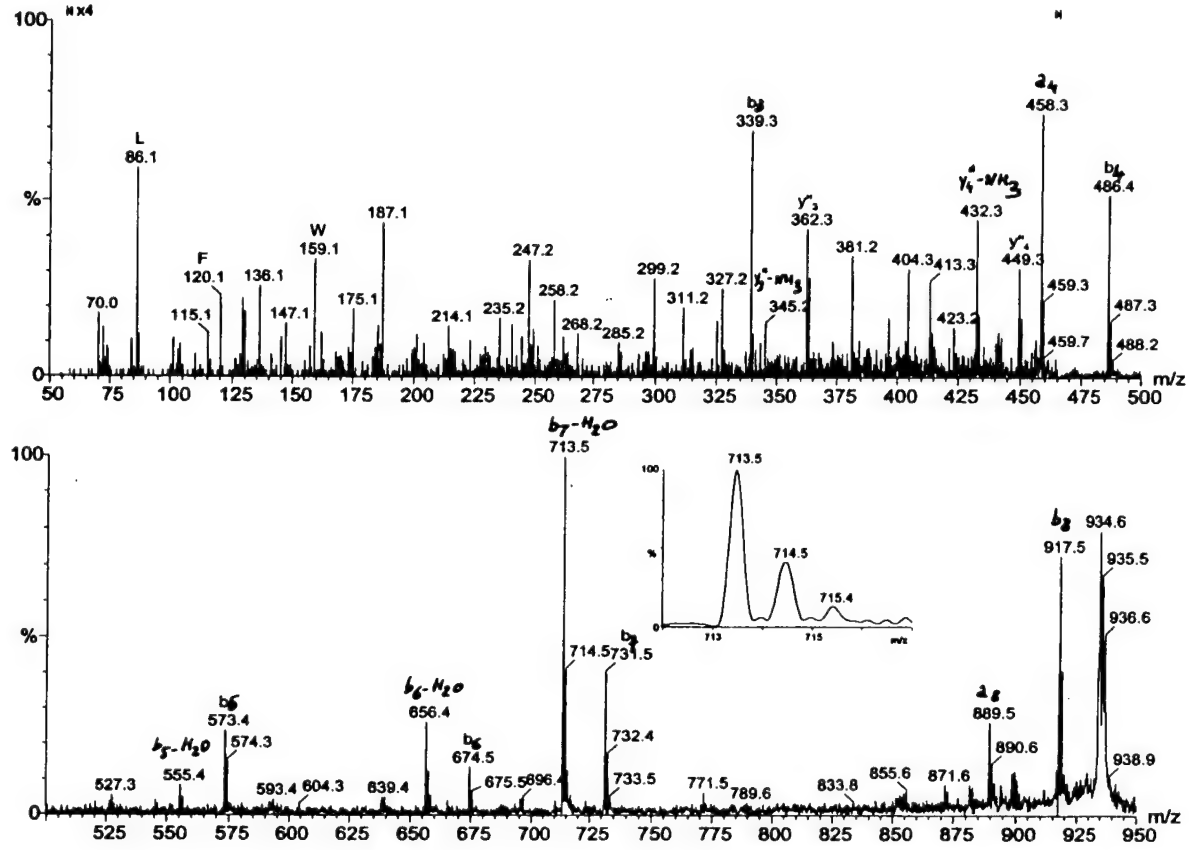


Figure 2. Q-TOF MS/MS spectrum of m/z 934.5, one of the smaller peaks in the locust Q-Tof MS spectrum (Fig. 1a). Sequence ions allowing unambiguous identification of the parent ion as Schistocerca AKH-II are annotated. Insert shows beautiful isotopic resolution of one of the product ions.

The MS system used was a Q-Tof (Micromass UK Ltd., Manchester,UK.), a hybrid quadrupole -time of flight tandem instrument [11]. In the present experiments the system was fitted with a nanoflow electrospray ion source. In this apparatus both MS and MS/MS spectra are acquired using the non-dispersive time of flight (TOF) analyser which has several advantages relative to a triple quadrupole machine (as used e.g. in [4]). By collecting all masses in the spectrum in parallel, rather than by scanning through the mass spectrum as a quadrupole does, TOF analysers record ion spectra with much greater efficiency. This allows data to be acquired with less sample being consumed and in a shorter timescale. At the same time the mass resolving power of the system is considerably higher than that obtainable from quadrupoles leading to significantly better signal to noise ratios in the spectrum and more reliable mass measurements (peaks are tall and narrow rather than short and wide).

In the MS mode the quadrupole of Q-Tof is operated in RF only mode, as a wide bandpass ion transport lens. No gas is admitted to the collision cell and ions are transmitted directly from the

source to the extraction cell of the orthogonal acceleration TOF device. A two stage extraction field accelerates ions into the drift tube and these ions are reflected by a single stage reflectron back to the two stage microchannel plate detector. Ion arrival times at the detector are measured by a 1GHz time to digital converter (TDC; Precision Instruments Inc., Knoxville, TN, USA) and stored in a [256 kBytes x 16 Bit] histogram memory continuously for the required integration period. This can be set as low as 0.1 sec or increased to several tens of seconds. In the data presented here the integration period was 5 sec. At the end of the integration period the spectrum is automatically written to disc and stored as a "scan" in the acquisition data file.

In the MS/MS mode of operation the quadrupole (MS #1) is scanned in conventional RF/DC mode and used to select the precursor ion for the MS/MS experiment. The precursor resolution can be varied and is usually adjusted to transmit about 3-4 Da to include isotope peaks. The collision cell is pressurized with Argon target gas and the collision energy raised to cause fragmentation of the precursor ion. The energy required for optimum fragmentation which is observed in the spectrum being acquired, and which depends on mass, charge state of the precursor and also on its molecular structure- is adjusted during the measurements. The operation and detection of ions in the TOF analyser (MS #2) is identical to the MS mode of operation.

Control of the instrument, acquisition of the data and subsequent post processing are carried out using the MassLynx PC based datasystem (Micromass UK Ltd., Manchester, UK).

Nanoflow electrospray ionization [12] using gold coated, drawn borosilicate glass needles with a tip orifice of about 1-2 µm produces very low sample flow rates (typically 10-50 nl/min) without reduction in sensitivity compared to the conventional mode (>5 µl/min) and in consequence large amounts of data may be acquired from as much as only one or a few microlitres of solution obtained from the sample preparation. After loading, the needle is fitted to the sample probe and inserted into the ion source to within about 7 mm of the atmospheric sampling cone. Applying a voltage (800-1600 V) is often sufficient to cause the electrospray process to spontaneously begin. Frequently, however, the application of additional pressure (0-2 bar) was beneficial in stabilising the spray. A small amount of nitrogen gas is flowed through the ion source to assist in desolvating ions produced by the nanoflow process. In the occasional event that the needle stopped spraying before being emptied, this was sometimes found to be associated with the tip becoming blocked. Removal of the needle from the ion source and breaking off a small part of the end of the tip under a laboratory microscope (x50 power) frequently got rid of the blockage and restored the spray allowing more data to be collected.

Like the insect corpus cardiacum, the model tissue in this study, all living cells, tissues, organs, organisms typically secrete -often species-specific- messenger molecules (in this case neuropeptides) by means of communication (an essential characteristic of "life" [13]). So do also bacteria [14;15;16]. Focusing on these types of communication molecules, in stead of on the bigger structural proteins/phospholipids, may be rewarding in an attempt to obtain fingerprint-style MS profiles for species identification. Q-Tof may be employed to help establish libraries of ABO MS profiles obtained from "crude" bacterial samples, completed with MS/MS data. All this should be done in the laboratory. In the field it may then be preferable to have robust portable equipment (small ESI-TOF systems or ESI ion traps [17]) which may be slightly less performant but able to generate data of sufficient quality and specificty to allow retrieval of the ABO from the Q-Tof database.

CONCLUSIONS

A hybrid nano-electrospray TOF apparatus as employed in this paper appears an ideal piece of equipment for use in characterization of ABOs. It can perfectly be used to work out what specifically needs to be detected in the field (and what sort of quantity is required). An important amount of effort will need to go in the selection of "easy preparable" analytes suitable to allow ABO identification. Databases will need to be set up for the various possible agents which need detecting. It is clear that the technology ideally needs to be transferred to field-portable systems. Small robust systems not requiring critical engineering tolerances (vibration etc.) such as provided by ion trap mass spectrometers may subsequently be appropriate devices for these purposes.

ACKNOWLEDGMENTS

The authors are endebted to Fund for Scientific Research-Flanders for financial support. Peter Verhaert is reasearch associate of the the same organization.

REFERENCES

1. Long SR (1997) Next-generation BW mass spectrometer for field detection. First Joint Services Workshop on Biological Mass Spectrometry, Abstract Volume: 13
2. Smith RD, Hofstadler SA, Udseth HR (1997) Front-end methodology for the rapid mass spectriometric characterization of biological agents. First Joint Services Workshop on Biological Mass Spectrometry, Abstract Volume: 47
3. Krishnamurthy T, Ross PL, Goode MT, Menking DL, Rajamani U, Heroux K (1997) Biomolecules and mass spectrometry. J. Natural Toxins 6: 121-162
4. Despeyroux D, Phillpotts R, Watts P (1996) Electrospray mass spectrometry for detection and characterization of purified cricket paralysis virus (CrPV). Rapid Comm Mass Spectrom 10: 937-941
5. Krishnamurthy T, Ross PL (1996) Rapid identification of bacteria by direct matrix-assisted laser desorption/ionization mass spectrometric analysis of whole cells. Rapid Comm Mass Spectrom 10: 1992-1996
6. Verhaert P, Vandesande F, De Loof A, Hoyes J, Bordoli R (1997) High speed sequences analysis of neuropeptides from single insects using nanoflow Q-TOF MS/MS. Proceedings of the 45th Congress of the American Society for Mass Spectrometrists. *in press*
7. Scarborough RM, Jamieson GC, Kalish F, Kramer SJ, McEnroe GA, Miller CA, Schooley DA (1984) Isolation and primary structure of two peptides with cardioacceleratory and hyperglycemic activity from the corpora cardiaca of *Periplaneta americana*. Proc Natl Acad Sci USA 81: 5575-5579
8. Witten JL, Schaffer MH, O'Shea M, Carter, Cook J, Hemling ME, Rinehart KL (1984) Structure of two cockroach neuropeptides assigned by fast atom bombardment mass spectrometry. Biochem Biophys Res Comm 124: 350-358
9. Stone JV, Mordue W, Batley KE, Morris HR (1976) Structure of locust adipokinetic hormone. Eur J Biochem 89: 195-202
10. Siegert K, Morgan P, Mordue W (1985) Primary structures of locust adipokinetic hormones II. Biol Cem Hoppe-Seyler 366: 723-727
11. Morris HR, Paxton T, Dell A, Langhorne J, Berg M, Bordoli RS, Hoyes J, Bateman RH (1996) High sensitivity collisionally-activated decomposition tandem mass spectrometry on a novel quadrupole/orthogonal-acceleration time-of-flight mass spectrometer. Rapid Comm Mass Spectrom 10: 889-896

12. Wilm M, Mann M (1996) Analytical properties of the nanoelectrospray ion source. Anal Chem 68: 1-8
13. De Loof A (1996) Wat is leven? Garant Press Leuven & Apeldoorn, 277 pp
14. Wirth R, Muscholl A, Wanner G (1996) The role of pheromones in bacterial interactions. Trends Microbiol 4: 96-103
15. Kaprelyants AS, Kell DB (1996) Do bacteria need to communicate with each other for growth? Trends Microbiol 4: 237-242
16. Gray KM (1997) Intercellular communication and group behavior in bacteria. Trends Microbiol 5: 184-188
17. Cooks RG, Wells MJ, Badman E (1997) Ion trap mass spectrometers: present performance and future capabilities. First Joint Services Workshop on Biological Mass Spectrometry, Abstract Volume: 7

Blank

MALDI VISUALIZATION OF THE *E. COLI* PROTEOME

Rachel R. Ogorzalek Loo, Charles Mitchell, and Philip C. Andrews
University of Michigan Medical School
Ann Arbor, MI 48109-0674

Ruth A. VanBogelen and Joseph A. Loo
Parke-Davis Pharmaceutical Research, Division of Warner-Lambert Company
Ann Arbor, MI 48105

ABSTRACT

A matrix assisted laser desorption ionization (MALDI)-based surface scanning technique recently introduced for measurement of intact masses of proteins directly from polyacrylamide gels was applied to whole cell extracts from *Escherichia coli*. Proteins from <3 kDa to > 120 kDa in size were detected at femtomole-level gel loadings. Individual mass spectra acquired at discrete pH increments on an isoelectric focusing gel can be consolidated into a *virtual* 2-D gel, (isoelectric point as the first dimension and MALDI-measured mass as the second). This parallel approach to analysis of gel-isolated proteins offers high throughput and is readily automated.

INTRODUCTION

High sensitivity, high throughput protein analysis is of intense interest today, particularly because it is a key technology for the post-genome world, where DNA information is tapped to identify and characterize proteins, which in turn can be used as markers for identification of organisms. Alterations in gene expression at the protein level reflect an organism's response to the environment, to disease, to drugs, or its developmental state. Modern technologies aim to probe these complex systems as a unit---as the sum of their proteins, rather than one protein at a time. Consequently, technologies based on parallel approaches are poised to contribute maximally. One parallel approach links two-dimensional polyacrylamide gel electrophoresis (2-D PAGE) to mass spectrometry (MS). 2-D gel electrophoresis is recognized widely for its ability to separate complex mixtures, resolving 1-2000 proteins on standard-sized, silver-stained gels, (with 75% of the proteins estimated to be present at levels below 500 fmol) and up to 10,000 proteins on large format gels (where 95% of the protein abundances are anticipated to lie below the limits of even high-sensitivity commercial Edman sequencers)[1]. The challenge, then, is to identify and characterize as many proteins as possible at as low a level as possible.

Most work linking mass spectrometry to polyacrylamide gel electrophoresis has concentrated on powerful peptide mapping approaches that identify proteins by database searching of the peptide masses or from sequences derived from the cleavage products.[2-4] These peptides must be eluted from gels or membranes prior to mass analysis. We have developed a matrix-assisted laser desorption/ionization (MALDI) MS-based surface scanning method in which proteins are desorbed *directly from polyacrylamide gels*, an approach attractive both for its simplicity and for its ability to provide masses of intact proteins at sub-picomole gel loadings[5-7], while avoiding the difficulties inherent in efficiently recovering proteins from gel matrices. An exciting aspect of this method is its potential contribution to high-throughput, high-sensitivity protein analysis. The method is also compatible with protein cleavage[5,6], although here we explore the utility of mass-analyzing *intact* proteins. Clearly the information obtained from analysis of intact proteins is complementary to that yielded by protein mapping, but it can be especially useful when there are questions about anomalous migration, about why the same protein has been identified at multiple locations on a gel, or about post-translational modifications when mapping approaches

have delivered low percent sequence coverages (such as when low abundance proteins are mapped from gels). Intact masses can also play a role in protein identification. For small genome organisms possessing minimal post-translational modifications and no introns (e.g., prokaryotes), identifications based only on a knowledge of pI to +/-0.05 pI units and an intact mass with 0.1% mass accuracy can uniquely identify a significant percentage of the theoretically expressed proteins and it dramatically narrows the number of potential proteins for the non-unique cases.[8]

In one approach, we take advantage of the orthogonality of mass spectrometry and isoelectric focusing (IEF) to create a *virtual 2-D gel*[5,7] from a one-dimensional separation. The term "virtual 2-D gel" is used for this IEF-MALDI-MS method, because isoelectric focusing comprises the first dimension of traditional 2-D gels, with mass as the second dimension. Even beyond the improved mass accuracy and mass resolution provided by the mass spectrometer, the virtual 2-D gel has several advantages over the traditional 2-D gel. Low molecular weight proteins (<10 kDa) usually migrate off the second dimension SDS gel, unless special electrophoresis conditions are employed. Moreover, the low molecular weight proteins are more difficult to stain as sensitively (on a pmol basis) as are larger proteins, because staining is particularly adept at detecting large proteins sensitively. These differences can have significant impact in studies characterizing the proteome, (*all* of the proteins potentially expressed by a genome). For example, 10% of the theoretical proteins predicted from the *E. coli* genome are below 12 kDa in mass. Low molecular weight proteins also arise from post-translational processing, and so an understanding of the *functional* genome requires characterization of low, as well as high molecular weight proteins. Finally, membrane proteins have been found to transfer poorly to SDS gels from the immobilized pH gradient gels often favored for IEF.[9] Virtual 2-D gels do not suffer from these limitations.

RESULTS AND DISCUSSION

Immobilized pH gradient (IPG) gels (pH 4-6.6) were employed for a one-dimensional IEF separation of an *E. coli* whole cell extract. These gels were employed because their superior gel-to-gel reproducibility enabled them to be mass-analyzed unstained, with little effort expended locating bands. Mass spectrometry performed on 20% of the pH range of a commercially available, precast IPG gel loaded with 10 µg of extract (quantitated by amino acid analysis) uncovered hundreds of proteins, many too faint to be visualized by silver-staining. Standard electrophoretic methods detailed by the manufacturer of the gels were employed. Figure 1 illustrates a silver-stained duplicate IEF gel loaded with the same quantity of protein as employed for the MALDI analyses. Individual mass spectra were acquired at approximately an 0.5 mm spacing spanning the pI 5.52-6.05 range of the gel (20% of its total span) from a gel soaked in sinapinic acid solution and allowed to air dry. Desorption was performed with 337 nm irradiation. Figure 2, assembled from the individual mass spectra acquired in this surface scan, illustrates that many large proteins are detected with this approach, even over this narrow pI range. Figure 3 presents an expanded view of this range.

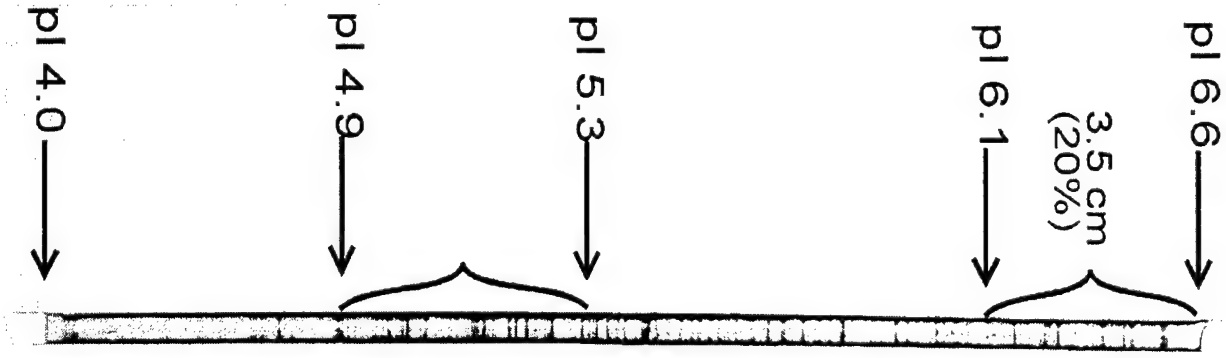
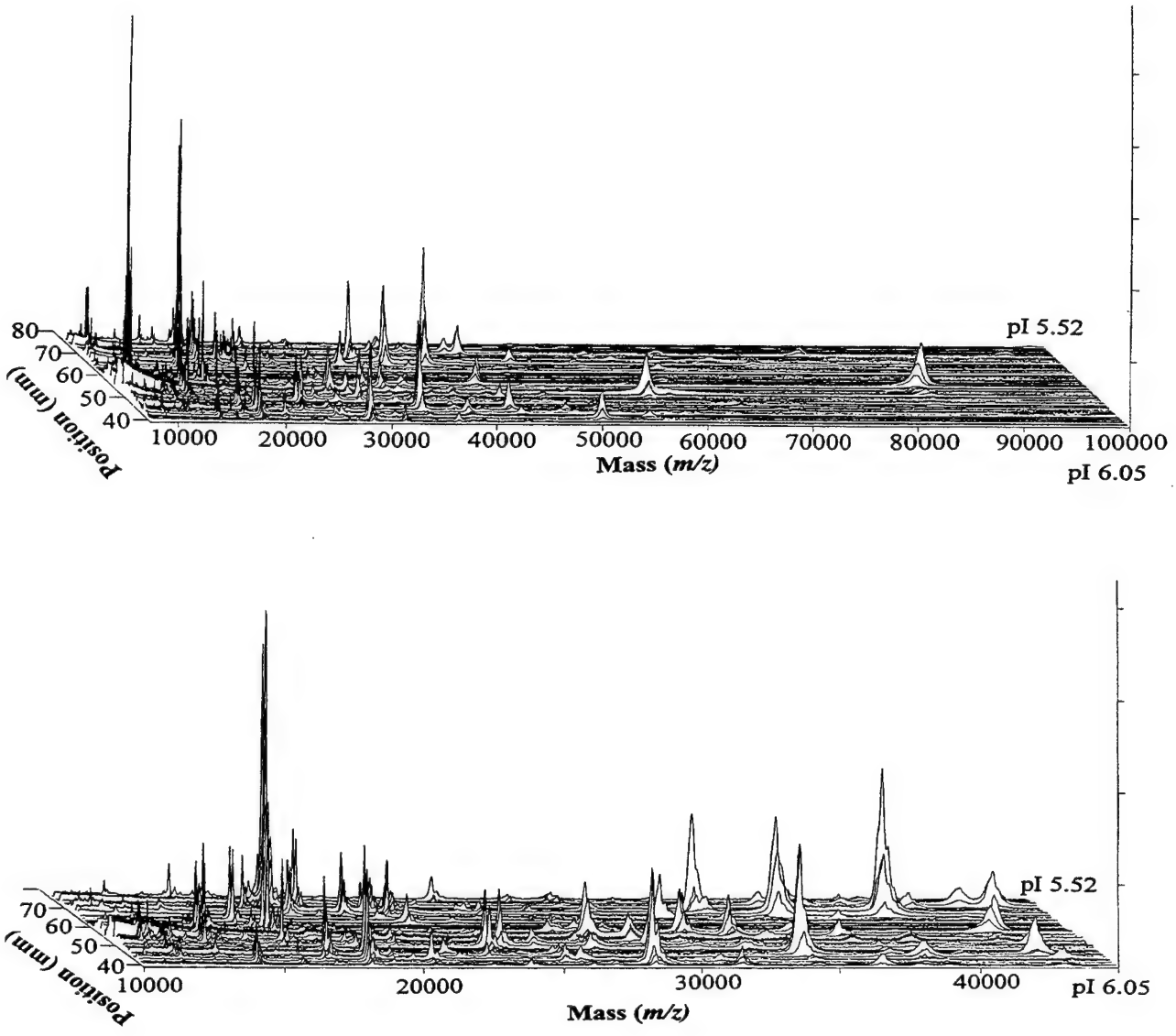


Figure 1. Duplicate immobilized pH gradient gel (silver stained).

There are 4285 unique proteins represented by the complete *E. coli* genome.[8] However, not all of these proteins may be expressed simultaneously; it is estimated that only a third would be expressed at any given time. Preliminary estimates of sensitivity indicate that proteins at abundances of less than 100 fmol loaded onto the gel are easily detected, and observed protein masses using this methodology ranged from below 3 kDa to greater than 120 kDa. In order to quantitate individual proteins based on initial sequencing yields, additional experiments were performed employing higher loadings of *E. coli* extracts onto IEF gels for subsequent blotting to membranes and Edman sequencing. These experiments indicated that *the most abundant* proteins detected by MALDI were present at loadings on the order of hundreds of femtomoles. Clearly, proteins present at 10 femtomole or even lower loadings on the gel are accessible to our methodology.



Figures 2-3. Plot assembled from individual MALDI-MS scans at 0.5 mm positions along a 1-D polyacrylamide gel, pI 6.05-5.52. Fig. 2 is an expansion.

Over the pI range 6.1-6.6, 32 bands were visualized by silver staining on the Fig. 1 IPG gel. Because there is considerable protein overlap in the one-dimensional separation, the number of bands stained on the IPG gel is lower than a true 2-D gel. When a duplicate IPG was transferred to a second dimension SDS gel and the resulting 2-D gel was silver stained, 123 proteins were detected over this region. MALDI-MS, however, detected 219 proteins (after discounting ions attributed to multiple charging). This corresponds to 50% of the 443 proteins predicted from the genome sequence to lie in this pI range (an artificially high number, given that all of the theoretically predicted proteins are not expected to be expressed simultaneously) and demonstrates that, on average, MALDI-MS visualization can be more sensitive than silver staining. With this methodology applied to small genome organisms, proteins can be identified strictly by accurate molecular weight and pI. Their identities can be confirmed by *in situ* CNBr cleavage.

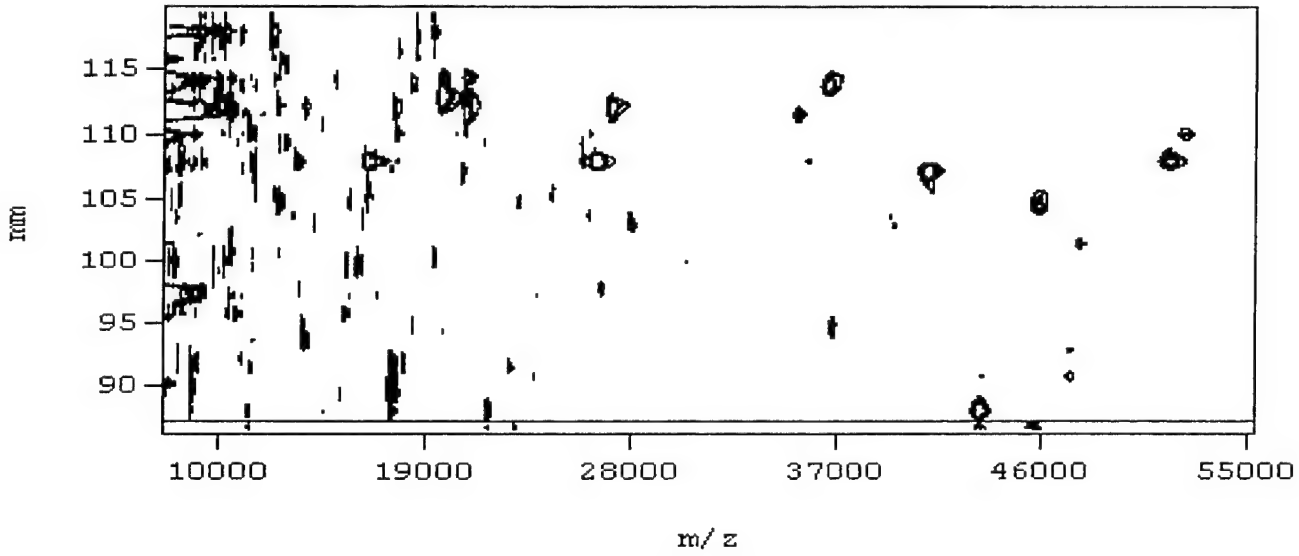


Figure 4. Virtual 2-D gel of 10 µgrams of *E. coli* whole cell extract. The pI range shown is 4.9-5.3 with the basic pI at the bottom of the image. The image was assembled from individual mass spectra acquired at a spacing of approximately 0.5 mm on a 3.5 cm length of an IPG gel.

Figure 4 illustrates an expanded view of a virtual 2-D gel spanning 4.9-5.3 pI units. In comparing this gel to classic 2-D gels, some differences need to be considered. The virtual 2-D gel presents a linear molecular weight scale, while a classic 2-D gel is logarithmic, condensing the higher mass spectral region. Our software has not yet discarded the spots due to multiply charged ions, so the virtual gel also has a few extra spots representing doubly charged highly abundant proteins. As mentioned earlier, staining usually is biased towards heavier proteins because it reflects nanogram rather than femtomole quantities of proteins. MALDI has the opposite bias and yields more intense signals for lower molecular weight proteins.

This approach is also useful for characterizing post-translational modifications--in the example illustrated in Fig. 5, the heterogeneity associated with Braun's lipoprotein, a 7 kDa membrane protein, is readily apparent from the mass spectrum of the intact protein desorbed from the gel.

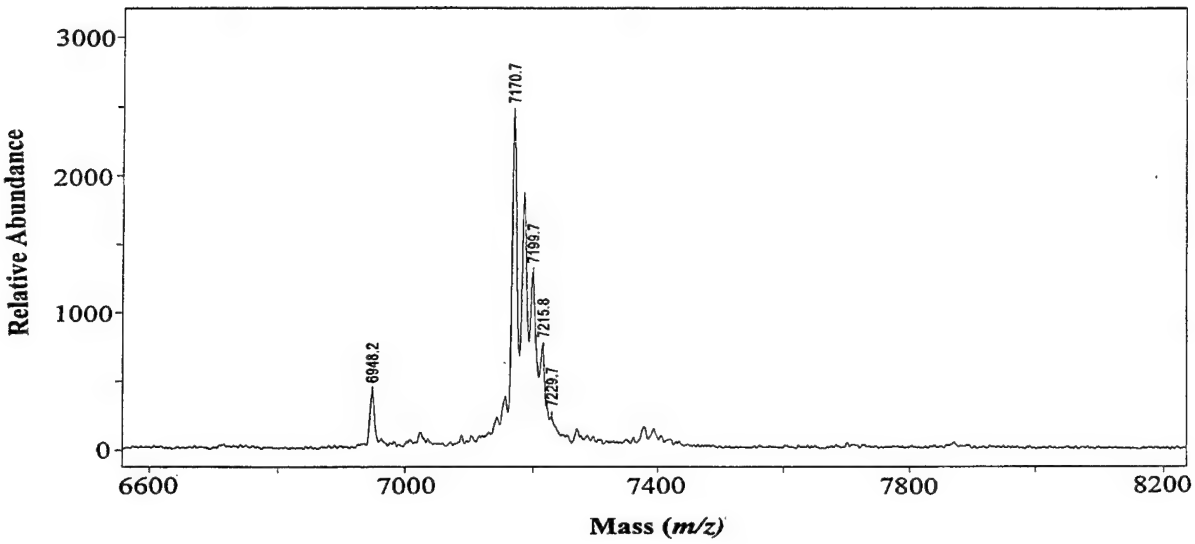


Figure 5. Detecting the lipid heterogeneity in Braun's lipoprotein.

CONCLUSIONS

Directly linking mass spectrometry to isoelectric focusing gels yields a virtual 2-D gel--the analogue to a classical 2-D gel, with MALDI mass spectrometry replacing the SDS electrophoresis second dimension. This combination is particularly advantageous for its ability to detect low molecular weight proteins prone to migration off the SDS dimension. IEF-MALDI-MS visualization was more sensitive than silver staining on average, and proteins from <3 kDa to >120 kDa were detected at femtomole-level gel loadings. Proteins can be identified strictly from pI and molecular weight for small genome organisms such as prokaryotes, once their genomes have been completely sequenced. This high-throughput approach sensitively probes entire organisms, either for identification or in order to examine their response to perturbations in their environment.

ACKNOWLEDGMENTS

R. R. O. L. thanks Parke-Davis Pharmaceutical for partial support through its post-doctoral fellowship program.

REFERENCES

[1] P. James, Biochem. Biophys. Res. Commun., 231 (1997) 1.
[2] S.D. Patterson and R. Aebersold, Electrophoresis, 16 (1995) 1791.
[3] M. Wilm, A. Shevchenko, T. Houthaeve, S. Breit, L. Schweigerer, T. Fotsis and M. Mann, Nature, 379 (1996) 466.
[4] A. Shevchenko, O.N. Jensen, A.V. Podtelejnikov, F. Sagliocco, M. Wilm, O. Vorm, P. Mortese, A. Shevchenko, H. Boucherie and M. Mann, Proc. Natl. Acad. Sci. USA, 93 (1996) 14440.
[5] R.R. Ogorzalek Loo, C. Mitchell, T. Stevenson, J.A. Loo and P.C. Andrews, in Techniques in Protein Chemistry VII, D. R. Marshak, Ed., Academic Press, San Diego, CA, 1996, p. 305.
[6] R.R. Ogorzalek Loo, T.I. Stevenson, C. Mitchell, J.A. Loo and P.C. Andrews, Anal. Chem., 68 (1996) 1910.

[7] R.R. Ogorzalek Loo, C. Mitchell, T.I. Stevenson, S.A. Martin, W. Hines, P. Juhasz, D. Patterson, J. Peltier, J.A. Loo and P.C. Andrews, Electrophoresis, 18 (1997) 382.
[8] J.D. Cavalcoli, R.A. VanBogelen, P.C. Andrews and B. Moldover, Electrophoresis, (in press) .
[9] C. Adessi, C. Miege, C. Albrieux and T. Rabilloud, Electrophoresis, 18 (1997) 127.

LC-MS/MS AND CE-MS/MS: MICRO-ANALYTICAL TOOLS TO STUDY BIOLOGICAL FUNCTION.

Ruedi Aebersold, Axel Ducret and Daniel Figeys
Department of Molecular Biotechnology, University of Washington, Seattle, WA 98195, USA

Powerful new technologies now permits the generation of complete gene expression profiles of specific cell types. Such profiles are dependent on and therefore diagnostic for the physiological or pathological state of the cell. Since most biochemical processes are mediated by proteins it is essential for the comprehensive analysis of biological systems that the level of expression and the state of activity of the proteins are also systematically investigated. Here we describe recently developed techniques for the establishment of protein expression profiles in cells and the functional investigation of biological systems and pathways. The approach is based on high resolution two-dimensional gel electrophoresis for the separation of complex protein mixtures and on electrospray ionization tandem mass spectrometry for the primary structural analysis of the separated proteins at high sensitivity.

INTRODUCTION

During the past two years, the complete genome sequences of a few selected species, among them the first genome sequence of a eukaryotic species, *Saccharomyces cerevisiae,* were completed and published. This represents a significant milestone in the history of biological science, since for the first time, the complete inherited information of biological species and with that, in principle, the composition of the components which constitute an organism have been determined. The comparative analysis of complete genome sequences encoded in the A,T,C,G four letter code is providing new evolutionary insights into genome organization and dynamics and has revealed a set of genes which appears to be close to the minimal number and composition required for a free living organism (1). However, the interpretation of the information encoded in the genome sequence in terms of the mechanism of biological processes has proven significantly more difficult.

Parameters describing the mechanism and the state of activity of a biological system are more likely to be deduced from the comprehensive quantitative comparison of samples representing two states of activity than from the analysis of the identity and functional state of individual molecules. On the mRNA level such comparisons are now feasible by means of a variety of recently described techniques (2,3). On the protein level the most advanced technique for establishing quantitative protein expression profiles is high resolution two-dimensional gel electrophoresis (2DE). While more technically limited than comparative analysis of mRNA samples, 2DE focuses on proteins, the biological effector molecules *per se.* The technique is well suited to detect indicators of the state of activity of both, biological systems and individual proteins.

In this report we describe 2DE as a tool for the analysis of regulated biological systems. Recent, dramatic technological advances in our ability to identify and characterize gel separated proteins moved 2DE beyond the stage of a purely descriptive technique. As a consequence of these developments 2DE and peptide mass spectrometry (MS) have converged into a powerful integrated analytical platform for the functional analysis of biological systems and pathways.

RESULTS AND DISCUSSION

2D Electrophoresis: A tool for probing biological function.

2DE was first shown to be a powerful technique for separating thousands of proteins in 1975 when O'Farrell and Klose independently developed a system that combined isoelectric focusing gel electrophoresis in the first dimension with SDS gel electrophoresis in the second dimension to create a pI by molecular weight separation matrix [4,5]. Using a large gel format, as many as 7000 protein spots could be separated [6]. It was soon recognized that quantitation of separated spots would allow the construction of a protein profile of a sample. Subsequent subtractive analysis of such profiles from samples generated under different experimental conditions would result in the detection of those proteins which were altered in a particular process within the background of proteins unaffected by the treatment. Furthermore, many such protein profiles could be archived in a computer for later comparative analyses.

For the functional characterization of biological systems one of the most significant attributes of 2DE is the multi dimensionality of data which can be extracted by the comparative analysis of suitable protein gel patterns. In addition to the two dimensions isoelectric point and molecular size which are used to separate the proteins into the 2D pattern and to give the technique its name, a number of additional parameters can be determined. They include the quantity of protein present in a spot, the type and stoichiometry of post translational modifications, the state of association of a protein with other proteins and its subcellular location. The determination of these parameters is significant because they are indicators of the state of activity of a biological system. In addition, the parameters can be followed over the course of time, thus revealing the dynamics of response to induced pertubations. 2DE therefore permits assessment of the state of activity of a system by the concurrent analysis of hundreds to thousands of proteins.

In the past, the use of 2D technology for the analysis of biological systems has been limited by technical difficulties in identifying separated and detected proteins. Without the possibility for rapid and conclusive protein identification 2DE is essentially a descriptive technique.

2DE and mass spectrometry: From descriptive pattern analysis to integrated protein analytical platform.

The ability to determine the identity of gel separated proteins dramatically enhances the power of 2DE for the analysis of biological systems. The most conclusive identifier for a protein is the amino acid sequence. Knowledge of the amino acid sequence of a protein not only provides unambiguous identification but is also key for further experimentation, including genetic manipulation of the sequence and expression levels of the targeted protein in experimental cell and animal systems. Identification of proteins separated by their amino acid sequence became routine with the introduction of suitable N-terminal and internal sequencing techniques (reviewed in 7). While accurate and general, chemical sequencing of gel separated proteins was relatively slow and insensitive. To enhance the sensitivity and the throughput of protein identification, a number of parameters which are characteristic for a particular polypeptide sequence have been used to identify proteins by correlating experimental data with objects in sequence databases. Such parameters include the amino acid composition, partial amino acid sequences, the masses of peptide fragments derived from a homogeneous protein sample and fragment ion spectra induced by collision-induced dissociation (CID) in a tandem mass spectrometer (reviewed in 8). Among these, CID spectra are preferred because i) the information gained from one peptide, in principle, is sufficient for the identification of a protein, ii) powerful software systems have been developed for the interactive [9,10] or automated [11,12] analysis of the spectra, iii) peptide sequences can be derived *de novo* in case the protein sequence is not contained in a sequence database [13] and iv) the method is fast and sensitive. It is therefore apparent that peptide (MS) and tandem mass spectrometry (MS/MS) have become the methods of choice for the identification of proteins separated by gel electrophoresis.

We and others have developed techniques for chemical and enzymatic protein fragmentation of gel separated proteins and the isolation and analysis by MS of the generated peptides (Reviewed in 4). The general procedure we have been using and which will be used to describe the experiments throughout this report is illustrated in Fig. 1. Proteins separated by gel electrophoresis are electroblotted onto nitrocellulose membranes and detected by staining. Individual spots or bands are excised and the absorbed proteins are digested with proteolytic enzymes on the membrane. The resulting peptide fragments are released into the supernatant and directly applied to one of the analytical system described in the following sections. Recently, variations on this basic procedure have been published which are too

numerous to list. They all have in common that selected proteins are subjected to proteolysis and that the peptide fragments are extracted for subsequent analysis, typically by MS. The procedures differ in the methods used to detect and isolate the intact protein, the conditions of proteolysis (in gel digestion, digestion after electroblotting on a membrane or digestion in solution) and the conditions used for extracting and potentially purifying the extracted samples prior to application to the MS, and the type of MS used for peptide analysis.

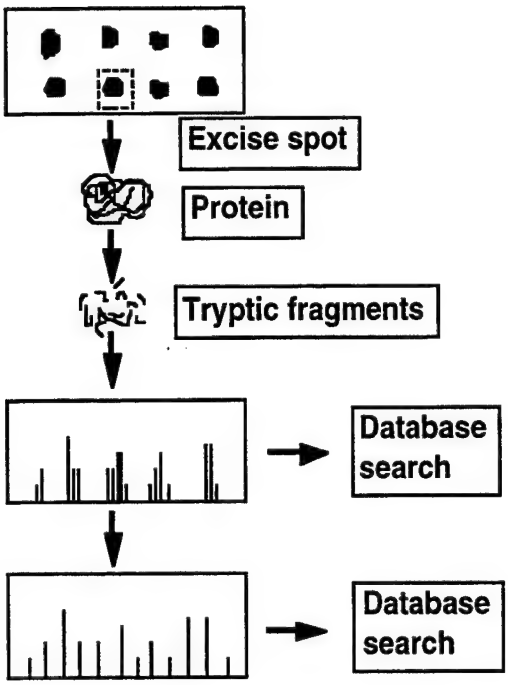


Fig. 1. Flow diagram of general procedure for identification of gel separated proteins.

Maximizing automation and throughput: LC-MS/MS.

Comparative analysis of 2DE patterns may highlight a sizable number of proteins which need to be identified in the course of functional analysis of a biological system. It has also been suggested that there is merit in the systematic identification of all the proteins in a biological sample [14]. Such projects, typically referred to as proteome analysis, require the unambiguous identification of hundreds to thousands of proteins. We therefore devised an integrated system for the automated, rapid and conclusive identification of proteins based on MS determination of polypeptide sequences [15]. The protein workstation consists of an autosampler, a microbore reverse-phase high performance liquid chromatography (RP-HPLC) system and ESI-MS/MS on-line detection and analysis of the eluting peptides. Protein digests are loaded into the autosampler and consecutively injected onto the RP-HPLC column. The separated peptides are then analyzed by MS/MS. If the ion current reaches a threshold pre-determined by the user, peptide ions are automatically subjected to CID. At the end of each LC–MS/MS run, the MS/MS spectra obtained are automatically analyzed by Sequest, a program that correlates the structural information of a CID spectrum with amino acid sequences contained in a database. Designed by John Yates and collaborators at the University of Washington in Seattle [11,12] , Sequest attempts to match an experimental MS/MS spectrum with the theoretical fragmentation pattern of all amino acid sequences contained in a database that are isobaric with the parent ion. This method does not presume any specificity of the enzyme used to generate the peptide and, because there is no formal interpretation of the MS/MS spectrum, the program can be run batchwise in a fully automated manner while the system acquires data from the next sample. Since with this method the CID analysis of an individual peptide is used to establish a database match, protein mixtures are easily characterized by the independent analysis of peptides originating from different proteins and the identification of a particular protein is usually based on the cumulative score of a number of peptides and therefore highly significant.

We have used this workstation successfully for the identification of 90 yeast proteins separated by 2DE in a completely automated manner. Protein spots containing an estimated 50–1000 ng of protein were pooled from three blots, digested with trypsin, and analyzed in two batches of 40 samples each (10 samples of the second batch were pools of at least two proteins which were co-digested) in unattended overnight or week-end runs. The samples were consecutively injected at a pace of one sample per 32 minutes by the autosampler, separated by RP-HPLC, and analyzed by MS/MS. At the end of each LC–MS/MS run, Sequest was automatically launched in the background as a post-acquisition application.

Proteins present in the gel at a level exceeding 200 ng were generally identified without difficulty by this strategy, while peptide maps generated by the digestion of proteins of lesser abundance were usually not sufficiently discriminating for identification purposes. Because the MS/MS spectrum of each peptide represents an independent analysis, unambiguous protein characterization was achieved, even when mixtures were present in a 2D gel spot. Therefore, LC-MS/MS analysis of protein digests represents a universal method for protein identification. In an automated mode, the protein workstation described here has proven to be a reasonably sensitive and extremely robust system.

Maximizing sensitivity: CE-MS/MS.

In contrast to DNA which can be readily amplified, protein analysis is limited to the samples that can be isolated from a biological source. A high level of sensitivity is therefore essential for the study of proteins, in particular regulatory proteins which are frequently of low abundance. We have developed an approach based on capillary electrophoresis (CE) MS/MS for the identification of proteins at very high sensitivity. The CE and chromatography-based technique described above share procedures for peptide sample preparation, automated generation of CID spectra in the ESI-MS/MS instrument and the algorithms used for protein identification. The two techniques differ by the peptide separation system that is integrated into the process. Recent developments in the field of MS have greatly increased the sensitivity achievable for protein and peptide analysis. For ESI-MS major sensitivity improvements were obtained by the introduction of microspray and nanospray ion sources [9,16,17]. Such ESI ion sources require the nl/min flow rates which are readily achieved by electro-osmotic flow in a fused silica capillary. We and others have previously demonstrated the power of CE combined with micro ESI-MS/MS [18,19,20], in particular for the rapid identification of proteins. Specifically, we have demonstrated a sensitivity of detection of less than 300 amol for peptides and the ability to generate CID spectra of peptides of sufficient quality for automated database searching with less than 600 amol of peptide applied to the system [18].

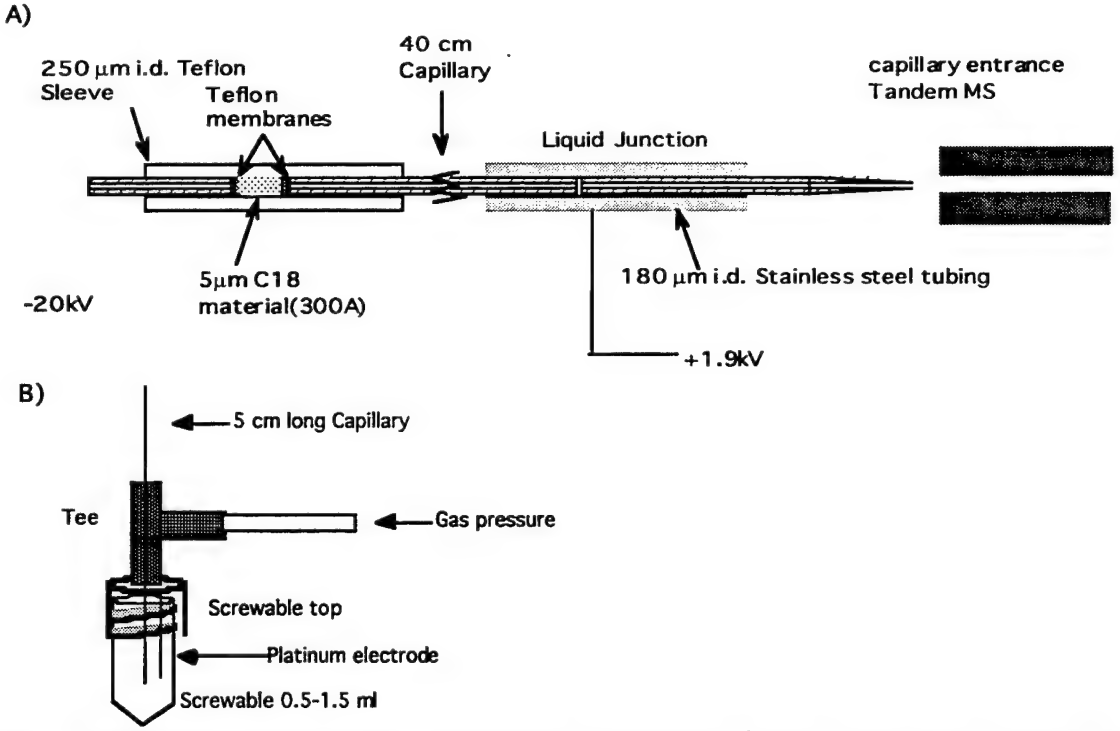


Fig. 2. SPE-CE-MS/MS system. Schematic illustration of SPE-CE-MS/MS device (A) and sample application system (B) are shown. Operation and performance of the system are described in the text.

By introducing an on-line solid-phase extraction (SPE) device, we were able to concentrate the analytes contained in large sample volumes into the small volumes compatible with CE. This eliminated the major limitation inherent in CE, namely incompatibility with the application of large sample volumes. The SPE device consists of a small RP chromatography column within the capillary which concentrates the analytes by reversible adsorption [19]. The system is schematically illustrated in Fig. 2.

Essentially, the protein of interest, typically separated by gel electrophoresis, is enzymatically cleaved off-line. Sample volumes of 10 to 30 µl of are typically obtained. The whole mixture of peptide fragments is injected onto the system and the peptides are accumulated on the SPE device. Once the system has been washed with an appropriate buffer, the peptides are eluted off the SPE device into the CE system and the CE experiment is started by applying a high voltage from the injection end of the capillary to the microelectrospray interface consisting of a liquid junction. The eluted peptides are electrophoretically stacked and separated. Using this approach, we have obtained concentration effects of up to 1000-fold. The peptides migrating out of the capillary are electrosprayed into the MS and analyzed as described above for the LC-MS/MS system. If the protein sequence is not represented in a database, *de novo* sequencing can be performed on the CID spectra. With this system we have obtained an absolute limit of detection (defined as the amount of analyte required to generate a signal size equal to 3 times the standard deviation of the background) of 660 amol in the MS mode and of 6 fmol in MS/MS mode and a concentration limit of detection (defined as the analyte concentration required to generate a signal equal to three times the standard deviation of the background signal if a volume of 20 ml of analyte was applied) of 300 amol/µl [19]. These detection limits clearly indicate that proteins can be identified with only a few fmol of material present in the digest. The system was used for the successful identification of low nanogram amounts of proteins separated by high resolution 2DE.

Integrating sample manipulation and analysis: Micro-machined analytical devices.

One of the challenges in high sensitivity protein analysis is the manipulation of protein and peptide solutions. The losses associated with sequential sample handling steps can be minimized if the surface wetted by the peptide sample is minimized. We have used microfabrication techniques to construct a device for the direct infusion of peptide samples into an electrospray ionization mass spectrometer. In this device (Fig. 3) etched sample and buffer reservoirs are connected *via* etched channels to a micro electrospray ion source. Peptide samples, typically unseparated tryptic digests of a protein, are applied to the different reservoirs. A flow of liquid originating in a specific reservoir is generated and selectively directed towards the microsprayer and the MS by electroosmotic pumping. The analyte proteins are identified by searching sequence databases with the information contained in the CID spectra of selected peptides. With this system we have achieved a limit of detection in the low fmol/µl range. We have also shown that samples deposited in different reservoirs can be sequentially mobilized without cross-contamination [20]. The direct coupling *on-line* of this microfabricated device to ESI-MS/MM represents an essential step towards the construction of an automated, high throughput, high sensitivity protein identification system.

CONCLUSIONS

The identification of proteins by their amino acid sequence has been a common and important process in diverse areas of biological research. The recent dramatic growth of large-scale genomic and expressed sequence tag (EST) sequence databases and their ease of accessibility through the Internet have created new challenges and opportunities for protein identification technology. The challenges are posed by the need to assign biological function to the numerous sequences representing potentially expressed proteins and to explain how the activities of individual proteins constitute a biological pathway or system. The opportunities are generated by the possibility for rapid and conclusive protein identification by correlating the information contained in their amino acid sequence with an object in a sequence database.

The analytical techniques described above and similar techniques developed in other groups now permit the identification of essentially any protein which can be detected by conventional staining methods. 2DE provides a convenient interface between the biological source of proteins and mass spectrometry based analytical techniques and permits the assignment of functional parameters to the separated proteins. Combined, 2DE and protein/peptide MS currently represent the most sensitive, comprehensive and quantitative system for the functional analysis of biological systems on the protein level. Future developments are likely to be directed towards the complete characterization of the covalent structure, i.e. the amino acid sequence as well as state of modification of proteins representing a specific state of activity. The experimental platforms and approaches described above are well suited for

this task. Rapid progress in the technology required for the structural analysis of post translationally modified proteins can therefore be expected.

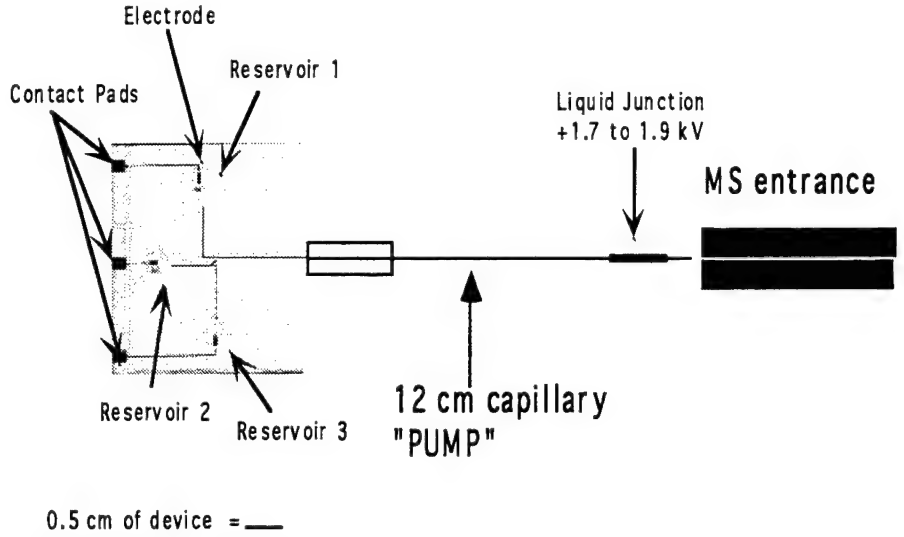


Fig.3. Schematic illustration of micro-machined analytical system. The channels on the device were graphically enhanced to make them more visible. Operation and performance of the device are described in the text.

ACKNOWLEDGEMENTS

Work was supported by the National Science Foundation (USA) Science and Technology Center for Molecular Biotechnology. DF was supported by consecutive fellowships from NSERC, Canada and from Zymogenetics Inc., Seattle WA.

ABBREVIATIONS

2DE: two dimensional polyacrylamide gel electrophoresis; CID: collision induced dissociation; MS: mass spectrometer/try; MS/MS: tandem mass spectrometer/try; RP-HPLC: reverse-phase high performance liquid chromatography; ESI: electrospray ionization; CE: capillary electrophoresis; SPE: solid-phase extraction; EST: expressed sequence tag.

REFERENCES

(1) Fraser CM, Gocayne JD, White O, Adams MD, Clayton RA, Fleischmann RD, Bult CJ, Kerlavage AD, Sutton D, Kelley JM et al. (1995) *Science* 270:397-403.

(2) Lennon G G and Lehrach H (1991) *Trends in Genetics*. 7:314-317.

(3) Schena M, Shalon D, Davis RW, and Brown PO (1995) *Science*. 270:467-470.

(4) O'Farrell PH (1975) *J. Biol. Chem.* 250 4007-4021.

(5) Klose, J (1975) *Humangenetik*. 26: 231-243.

(6) Klose, J (1989) *Electrophoresis.* 10:140-152

(7) Aebersold, R and Leavitt, J (1989) *Electrophoresis.* 11:517-527.

(8) Patterson SD and Aebersold R (1995) *Electrophoresis* 16:1791-1814.

(9) Mann M and Wilm M (1994) *Anal. Chem.* 66:4390-4399.

(10) Clauser KR, Hall SC, Smith DM, Webb JW, Andrews LE, Tran HM., Epstein LB, Burlingame AL (1995) *Proc. Natl. Acad. Sci. USA* 92:5072-5076.

(11) Eng J, McCormack AL and Yates JR. III., (1994) *J. Am .Soc. Mass Spectrom.* 5: 976-989.

(12) Yates JR. III, Eng JK, McCormack AL and Schieltz D (1995) *Anal. Chem.* 67:1426-1436.

(13) Wilm M, Shevchenko A, Houthaeve T, Breit S, Schweigerer L, Fotsis T and Mann M, (1996) *Nature* 379:466-469.

(14) Wilkins MR, Sanchez J-C, Gooley AA, Appel RD, Humphery-Smith I, Hochstrasser DF and Williams KL. (1996) *Biotechnol. Genet. Engin. Rev.* 13:19-50.

(15) Ducret A, van Oostveen I, Eng JK, Yates JR III and Aebersold, R. submitted

(16) Smith RD (1988) *Anal. Chem.* 60:1948-1953.

(17) Emmett M and Caprioli RM (1994) *J. Am. Soc. Mass. Spectrom.* 5:605-613.

(18) Figeys D, van Oostveen I, Ducret A and Aebersold R. (1996) *Anal. Chem.* 68:1822-1828.

(19) Figeys D, Ducret A, Yates JR III and Aebersold, R. (1996) *Nature Biotechnology.* 14:1579-1583.

(20) Figeys D, Ning Y and Aebersold R. (1997) *Anal. Chem.* in press.

Blank

FRONT -END METHODOLOGY FOR THE RAPID MASS SPECTROMETRIC CHARACTERIZATION OF BIOLOGICAL AGENTS

Steven A. Hofstadler, Chuanliang Liu, Scott A. Shaffer, and Richard D. Smith
Pacific Northwest National Laboratory
Richland, WA 99352

A. Peter Snyder
Edgewood Research, Development, and Engineering Center
Aberdeen Proving Ground, MD 21010-5423

ABSTRACT

The identification of potential biological agents in field situations currently presents a major analytical challenge, demanding high sensitivity, speed and the ability to distinguish biological agents in a manner that greatly minimizes the occurrences of both false positives and false negatives. Chemical characterization methods are attractive due to their potential speed, but pose significant technical demands in regard to sensitivity and the ability to deal with the complexity of biological samples and the large "environmental background". The most promising, rapid, flexible and broadly applicable potential technology for accomplishing this is mass spectrometry, but the ability to apply this technology is crucially dependent upon the "front-end" processing that is required to deliver a tractable sample to the analyzer. The development of rapid "front end" technology for sample processing, fractionation and clean-up in conjunction with the analysis capabilities provided by electrospray ionization with MS/MS analysis using ion trap mass spectrometers is required to effectively implement this technology. It is this aspect of the ESI-MS biological agent characterization methodology we describe in this work; the rapid automated on-line "front end" to the mass spectrometer. In this work we describe the development of a key isolation/fractionation step based on dual microdialysis which renders the sample amenable to high performance mass spectrometric analysis.

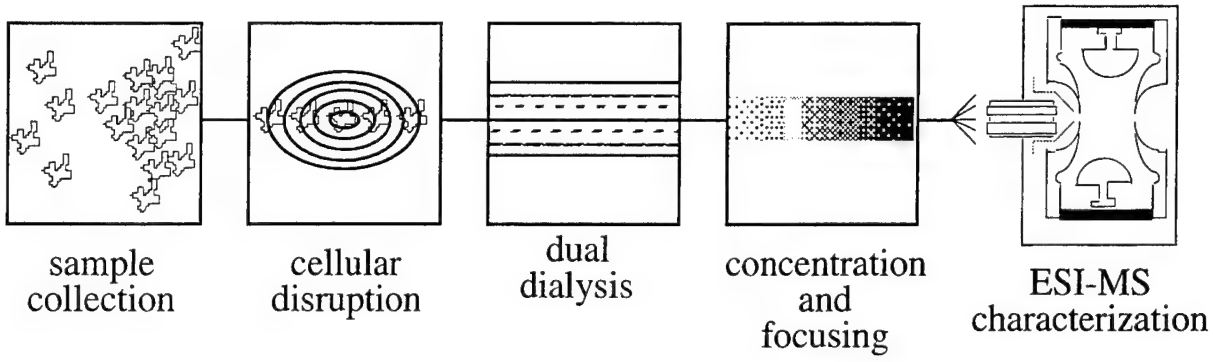
INTRODUCTION

To date there has been an enormous and continuing investment in the development of more compact, sensitive and powerful MS instrumentation, particularly that based around quadrupole ion trap instrumentation (e.g., for the CBMS as well as for field use in other applications). It is increasingly clear, however, that the potential for fielding sufficiently fast, sensitive and reliable automated instrumentation is not limited so much by the speed or sensitivity of the mass spectrometer, but rather by the ruggedness, speed and efficacy of sample processing, enrichment and purification schemes which are a necessary prelude to effective mass spectrometry based biological threat identification.

The use of electrospray ionization (ESI) with multi-stage mass spectrometry (i.e., MS/MS) is particularly well suited since it readily accommodates liquid sample processing and transfer. Additionally, the multiple charging obtained with ESI makes MS/MS approaches significantly more effective (due to the much high efficiency of collisional dissociation of highly charged ions compared to singly charged ions as typically produced by other ionization methods), and increases the ability to identify threat agents in complex environments, mixtures, and at lower levels. The sensitivity, gentleness, and broad applicability of ESI-MS is well documented and recent work has demonstrated particular promise for the on-line and automated characterization of microorganisms and biotoxins. Although matrix-assisted laser desorption ionization (MALDI) can conceptually also be coupled effectively with parallel sample processing schemes, the present research focuses on the use of ESI due to the direct compatibility provided with low flow rate liquid sample processing schemes and the fact that ESI produces higher charge states of large molecules [1-3], which affords more effective collisional dissociation for more sensitive or selective analyses.

Our goal is to develop and implement more effective and rapid methods for the analysis of complex biological samples, and specifically methods for the rapid analysis of microorganisms and other potential biological threats. The successful implementation of this approach would enable the

fielding of rugged and reliable instrumentation for the rapid identification of microorganisms based upon the detection of distinctive polypeptide biomarkers at very low levels. The potential sensitivity and selectivity of this approach arises directly from the combination of the front-end sample processing and the selective ion accumulation capability of ion trapping instrumentation. Rendering a sample amenable to ESI-MS characterization of smaller proteins involves the following steps: 1) collection, 2) cell lysing, 3) protein separation and/or sizing, 4) protein isolation and/or concentration, and 5) ESI-MS characterization of the targeted protein(s). The five steps are illustrated pictorially below.



This approach employs a novel on-line dual-dialysis approach, developed at out laboratory as a purification step to select a molecular weight fraction that includes small proteins, and the use of capillary isoelectric focusing (CIEF) for separation and concentration of selected proteins. The effluent from the CIEF separation will then be directly interfaced with ESI-MS for characterization of distinctive higher abundance and lower molecular weight proteins.

New on-line microdialysis methodologies for sample handling and processing

We have recently demonstrated new and versatile methods for nano-scale sample clean up based on an approach which employs a microdialysis membrane in conjunction with a counter-current flow of a suitable buffer [4-7]. While still in a developmental stage, this approach has already proven to be an effective method for rapid desalting of proteins and nucleic acids, an essential step for effective ESI-MS analysis. This capability has already significantly enhanced the ability to work with biological samples originating from complex sample matrices. Recent experiments also indicate that the dual microdialysis approach, in which membranes of different molecular weight cut-off (MWCO) are employed (see later discussion), has the potential to provide a rapid and effective means of selecting a molecular weight (MW) fraction of biomolecules. Of direct significance to this project is the potential for using microfabricated devices to carry out these steps on even smaller sample volumes, to do this more rapidly and in a fashion that allows direct ESI-MS analysis

On-line Microdialysis

Dialysis is a commonly used sample clean-up method. The new on-line microdialysis method is distinguished by its speed, on-line nature and the exceptional quality of the results obtained with extremely small samples. As one application, we have found this approach invaluable for ESI-MS analysis of PCR products, a problem that had previously proved intractable at our laboratory due to the lower efficiency or greater sample losses associated with standard clean-up [4, 7]. Microdialysis is also effective for removing other low molecular weight constituents from complex systems such as cell lysates, and provides an easy route for the rapid modification of buffer composition for subsequent microscale sample manipulations.

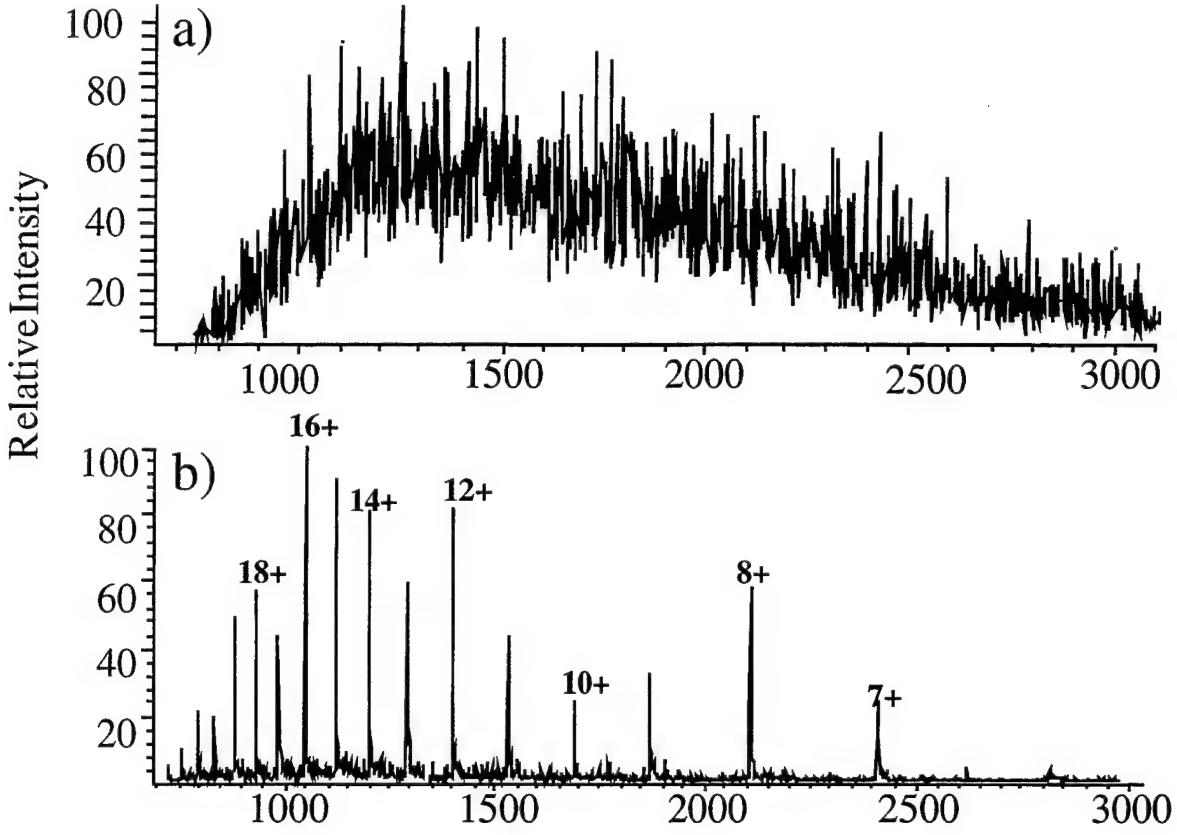


Figure 1. Micro-dialysis enables rapid removal of salt, buffering agents, and other intereferants from solution prior to ESI-MS characterization. a) spectrum acquired from a 3 μM solution of apomyoglobin (MW = 16,951 Da) in 250 mM NaCl, b) spectrum from 3 μM apomyoglobin in 1.0 M NaCl acquired following an on-line microdialysis step.

Our aim is to implement a dual microdialysis approach for the rapid clean-up and fraction selection of crude cellular lysates that is interfaced for effectively simultaneous mass spectrometric analysis. Inherent in this is the capability for sampling soluble polypeptides and smaller proteins amenable to ESI, while discarding insoluble membrane and constituents and other cellular residue. The goal of the dual microdialysis step is to combine the ability to discard undesired constituents simultaneously with target-component clean-up and desalting steps.

The conceptual implementation of this approach is relatively straightforward; the cell sample or crude lysate is dialyzed using membrane of high molecular weight cut off (MWCO) (e.g. 25,000 Da in our preliminary work). This dialysate is subsequently dialyzed with a low MWCO membrane (e.g. 1,000 Da) to remove low molecular weight interferants including environmental contaminants, buffering agents, and low molecular weight biomolecules. This separates the crude lysate into three roughly sized fractions: a fraction containing cellular debris and molecular species in excess of ≈ 25,000 Da, a fraction containing molecular species in the range of ≈ 1,000 - 25,000 Da, and a fraction containing low molecular weight species. The protein fraction containing species of the targeted molecular weight range (1,000 - 25,000 Da in this example) will be significantly depleted of species outside the molecular weight range of interest which would otherwise interfere with subsequent micro-column manipulations and ESI-MS characterization. Figure 2 illustrates the design used for our initial evaluation of the dual microdialysis approach.

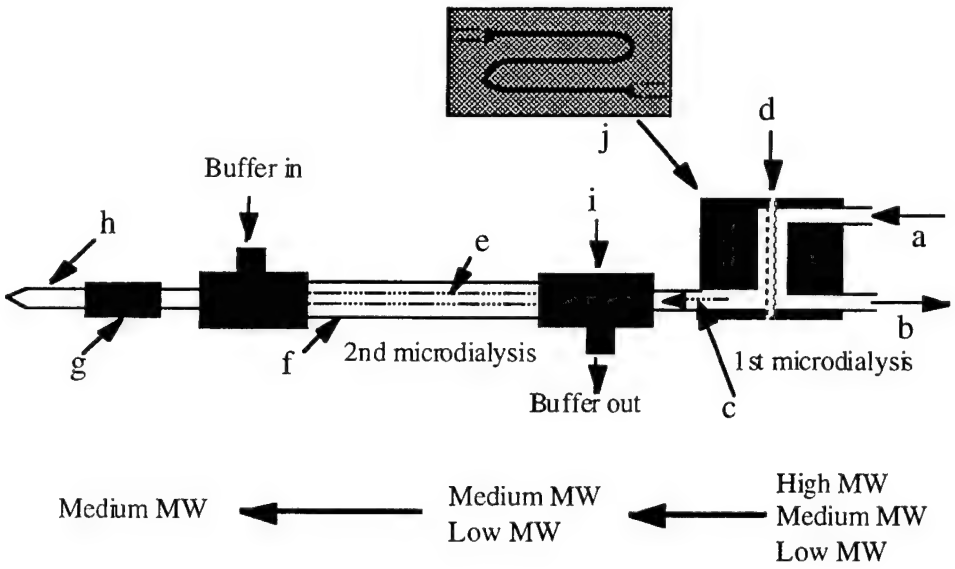


Figure 2. Schematic of the initial dual microdialysis arrangement. a, sample injection inlet; b, partial sample outlet; c. selected sample fraction outlet to second stage of microdialysis; d. MWCO 30K regenerated cellulose membrane for the first stage microdialysis; e. MWCO 13K regenerated cellulose microdialysis fiber for the second stage microdialysis; f. Teflon tube; g. low dead volume union used for on-line high voltage application; h. pulled 50 μm id fused silica capillary micro-ESI source; i. low dead volume "tee"; j. front-view of the serpentine channel of first stage.

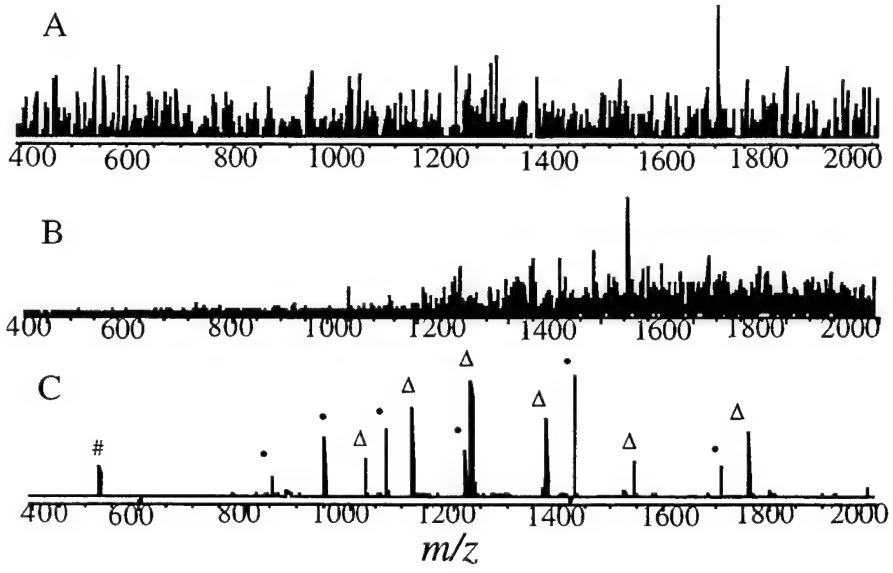


Figure 3. ESI-MS spectrum of a protein/polypeptide mixture consisting of BSA in large excess (2 mg/mL), 0.05 mg/mL cytochrome *c*, 0.02 mg/mL ubiquitin and 0.01 mg/mL bradykinin in 0.5 M NaCl. A) direct infusion; B) after a single stage microdialysis using a MWCO 13K membrane; C) after dual microdialysis. Note the effective removal of both BSA and NaCl is needed for detection of the 3 polypeptides (# bradykinin, • ubiquitin and Δ cytochrome *c*).

An initial demonstration of the dual microdialysis approach is shown in Figure 3. The results showed that both undesired high molecular weight constituents (albumin in large excess was used as a surrogate for these components) as well as other cellular residue could be effectively removed, which is essential since MS quality is otherwise greatly degraded. It is evident from the results shown that both stages of microdialysis are needed and that the present approach is effective in rapidly selecting a representative MW fraction amenable to ESI-MS analysis.

As shown in Figure 4, the dual microdialysis approach yielded an informative spectrum after the direct introduction of a cell lysate to the device. Due to the complex nature of the sample, the spectrum is filled with numerous peaks. Expanded views of a small *m/z* range (Figure 4, left) show the detailed mass spectrometric data obtainable. With the initial dual microdialysis configuration, the time needed for obtaining a useful spectrum from the crude cell lysate is ~20 minutes, a time that should be greatly reduced by implementation as a microfabricated device that would largely eliminate dead volumes and also accrue further benefits in speed and efficiency due to the reductions in channel dimension possible. (The smallest available microdialysis fiber has a diameter of ~200 μm, a factor that limits not only speed, but also the minimum flow rate needed.) An additional characteristic of the device has been its rugged performance, operating for more than a week of intensive use. We have determined a major factor in realizing this performance is the first stage design, where the sample flow effectively sweeps cellular residue away from the membrane, rather than rapidly clogging the membrane as would results for conventional ultra-filtration. Considering that most of the dead volume in the present design arises from the connecting tubing (the dead volume from the serpentine channels and microdialysis fiber combined is only ~20 μL), implementation as a microfabricated device should easily decrease the achievable analysis time to < 5 minutes.

Information from MS/MS results of selected cellular components will be very important for the characterization of target cellular compounds and identification of microorganisms. To explore the possibility of probing the structure of the individual cellular component, MS/MS was performed on selected ions in the spectrum obtained from the same cell lysate. Figure 4 (lower right) shows the MS/MS spectrum of the *m/z* 1355 species obtained from isolation and collision induced dissociation (CID). Abundant fragment ions were obtained which provide important "fingerprint" information for the selected ion.

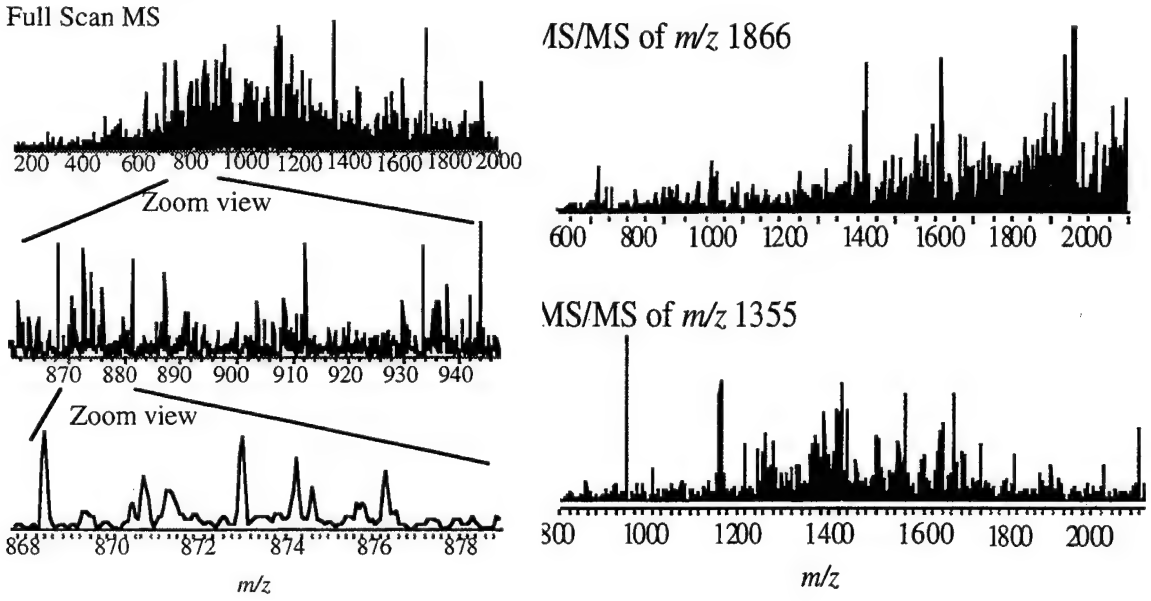


Figure 4. Direct ESI-MS analysis of bacterial cells following dual micro-dialysis, showing sensitive detection using a conventional ESI-ion trap mass spectrometer (left). Mass spectra from dissociation of two peaks from the spectrum on the left are shown on the right.

CONCLUSIONS

The dual-microdialysis approach has been shown to enable fast and efficient sample fractionation and clean-up for ESI-MS analysis of complex biological samples. The high molecular weight contaminants were effectively removed by passing through the first stage microdialysis and the low molecular weight interfering species were removed by the second stage microdialysis. The molecular weight range of samples of interest can be controlled by use of appropriate MWCO membranes at both stages. Direct ESI-MS analysis of a crude cell lysate after passing through the dual-microdialysis system produced a clean and informative spectrum. This approach should be generally applicable for ESI-MS analysis of any complex sample or in other applications demanding samples free of similar contaminants.

REFERENCES

1. Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F. and Whitehouse, C. M., *Mass Spectrom. Rev.*, **1990**, *9*, 37-70
2. Smith, R. D., Loo, J. A., Loo, R. R. O., Busman, M. and Udseth, H. R., *Mass Spectrom. Rev.*, **1991**, *10*, 359-451
3. Gaskell, S. J., *J. Mass Spectrom.*, **1997**, *32*, 677-688
4. Muddiman, D. C., Wunschel, D. S., Liu, C. L., Pasatolic, L., Fox, K. F., Fox, A., Anderson, G. A. and Smith, R. D., *Anal. Chem.*, **1996**, *68*, 3705-3712
5. Liu, C., Wu, Q., Harms, A. C. and Smith, R. D., *Anal. Chem.*, **1996**, *68*, 3295-3299
6. Wu, Q., Liu, C. L. and Smith, R. D., *Rapid Comm. Mass Spectrom.*, **1996**, *10*, 835-838
7. Liu, C. L., Wu, Q. Y., Harms, A. C. and Smith, R. D., *Anal. Chem.*, **1996**, *68*, 3295-3299

ACKNOWLEDGEMENTS

This research was supported by the US Army Edgewood Research Development and Engineering Center and through internal exploratory research and the Director, Office of Biological and Environmental Research, U.S. Department of Energy. Pacific Northwest National Laboratory is operated by Battelle Memorial Institute for the U.S. Department of Energy through contract DE-AC06-76RLO 1830

ION TRAP MASS SPECTROMETERS: PRESENT PERFORMANCE AND FUTURE CAPABILITIES

R. Graham Cooks, M. J. Wells and E. Badman
Department of Chemistry, Purdue University
West Lafayette, IN

ABSTRACT

The developments which led to the emergence of the ion trap as a sensitive, highly capable tandem mass spectrometer are reviewed. Resonant ejection, extension of the mass/charge range to very high values and injection of ions from external sources get special emphasis and the detection of peptides in trace quantities by these methods is illustrated. Further progress depends upon a more complete knowledge of the motion of trapped ions. This information is being sought by laser dc pulse tomography experiments and by multi-particle simulations. How this information is being used to improve analytical performance of the trap and to develop new methods of operation is illustrated. Special attention is given to three new experiments: (i) the use of fourier transform methods in conjunction with non-destructive detection, (ii) the characteristic of ion traps with cylindrical geometry and (iii) the use of affinity membrane introduction systems for rapid sampling of targeted compounds. It is suggested that these developments should lay the foundation for future miniaturized, portable ion trap mass spectrometers which retain the high performance of the present device.

Blank

DETERMINATION OF BACTERIAL PROTEIN PROFILES BY MATRIX-ASSISTED LASER DESORPTION/IONIZATION MASS SPECTROMETRY INTERFACED TO CAPILLARY SEPARATION METHODS

Teresa Cain, Xaoili Liang, Lin Liu and David M. Lubman
Department of Chemistry
The University of Michigan
Ann Arbor, Michigan

ABSTRACT

A rapid method for profiling bacterial and cellular proteins has been developed using a combination of capillary separations followed by (MALDI-MS) matrix-assisted laser desorption/ionization mass spectrometric analysis. In this method, bacteria are sonicated, the cell walls broken, and the water-soluble proteins precipitated for analysis. In initial work the proteins were separated by capillary liquid chromatography and detected on-line by a UV absorption detector. In more recent work protein bands have been separated and detected using capillary isoelectric focusing. The protein fractions are then collected for off-line analysis by MALDI/MS. Using this method, it is demonstrated that bacteria can be discriminated based upon their protein profiles to the species level with subpicomole level detection of proteins. More recent work describing attempts to solubilize and detect high mass fat soluble protein fractions as potential biomarkers will also be discussed.

Blank

MALDI-TOFMS for Direct Analysis of Bacteria Proteins: Effects of Protein Extraction and Sample/Matrix Preparation on Mass Spectral Patterns

Zhengping Wang, Larry Russon, and Liang Li*
Department of Chemistry, University of Alberta, Edmonton, Alberta, Canada T6G 2G2

Dennis C. Roser[1] and S. Randolph Long[2]*
[1]Gunpowder Branch, Aberdeen Proving Ground, MD 21010;
[2]Edgewood RDE Center, SCBRD-RTE, Aberdeen Proving Ground, MD 21010, USA

Abstract

Matrix-assisted laser desorption ionization (MALDI) time-of-flight mass spectrometry (TOFMS) can be used for rapid detection of bacterial proteins in a crude mixture. However, the performance of MALDI in direct analysis of protein mixtures strongly depends on the physical and chemical nature of the sample, the selection of appropriate matrices and solvents, and the sample/matrix preparation process. The objective of this work is to determine the issues involved in sample handling and sample/matrix preparation procedures for profiling bacterial peptides and proteins directly from cell suspensions or cell extracts. It is demonstrated that sample/matrix preparation and the method of protein extraction can affect the mass spectral patterns significantly. However, with the use of well-controlled experimental conditions, reproducible spectra can be readily obtained.

1. Introduction

Rapid and unambiguous identification of bacterial species and strains has a broad appeal to in-field hazard assessment as well as clinical and environmental applications. One possible approach to identifying bacteria is to examine the protein content and seek protein biomarkers that are unique to specific bacteria. The type of proteins present in bacterial cells is immense and their amounts vary greatly. Two-dimensional gel electrophoresis analysis of proteins has revealed that at least 1000 different proteins are present in Escherichia coli cell extracts [1]. One or several proteins among such a large pool of proteins may be used as the protein biomarkers for bacteria identification. The challenge is to identify these biomarkers. The selection of suitable biomarkers may well be dictated by the detection method used. The protein biomarkers identified should be readily analyzed and less prone to interference from other species in bacteria cells and from background molecules present in real world samples.

With the recent advances in ionization and detection methods, mass spectrometry has become a powerful tool for protein analysis. Both matrix-assisted laser desorption ionization (MALDI) and electrospray ionization (ESI) mass spectrometry can potentially be used as a means for rapid identification of bacteria. The advantage of the MALDI technique is its ability to analyze crude protein mixtures without the need of extensive separation, although on-line analysis is currently difficult. Several researchers have shown the possibility of using MALDI time-of-flight mass spectrometry for rapid identification of bacteria based on mass spectral patterns [2-4]. In these methods, proteins were first extracted from bacteria cells. The crude protein mixtures were then analyzed by MALDI-TOFMS. One or more peaks detected in the mass spectrum, which were found to be unique for a certain type of bacteria, were used for identification.

It is well known that the performance of MALDI in direct analysis of mixtures strongly depends on the physical and chemical natures of the sample, the selection of appropriate matrices and solvents, and the sample/matrix preparation process. To further advance the MALDI method for bacteria identification, there is a strong need to characterize the effects of protein extraction and sample/matrix preparation methods on mass spectral patterns. It is hoped that such a detailed characterization will result in the selection of optimal biomarkers, enhance the analytical

performance in terms of detection sensitivity, selectivity, and tolerance to interference, and improve the reproducibility of the technique.

In this report, the influences of matrix/sample preparation methods on the detectability of proteins from the bacteria cell extracts are examined. Some preliminary results on the effect of the cell extraction procedures on the mass spectral patterns are also presented.

2. Experimental

U of A Experiment. The bacteria samples studied in this work include Bacillus subtilis, Bacillus thuringiensis, Bacillus cereus, and Escherichia coli. About 1 to 1.5 mg of lyophilized bacteria were suspended in 250 µl solvent. The type of solvent tested for dissolving proteins from the cells include water, 0.1% trifluoroacetic acid (TFA), isopropanol, and a solvent mixture consisting of 17% formic acid, 33% methanol, and 50% water. The cell suspension was vortexed for about 2 min and then centrifuged. The supernatant solution was taken for MALDI analysis.

The two-layer method was used for matrix/sample preparation [5]. Both α-cyano-4-hydroxycinnamic acid (HCCA) and sinapinic acid were initially examined as the matrices for bacteria protein analysis and it was found that HCCA provided better sensitivity. Consequently, HCCA was used throughout this study. In the two-layer method, the first layer was formed by placing 1 µl of 100 mM HCCA in 99% acetone and 1% water (v/v) on the MALDI probe tip and allowing it to dry in air. A saturated solution of HCCA was prepared in a solvent mixture. A variety of solvent mixtures were used and their compositions are described in the Results and Discussion. The saturated solution was 1:1 by volume mixed with the sample solutions. 1 µl of this solution was placed on top of the first layer and allowed to dry. The tip was then dipped in pure water for approximately 10 s and the excess water was shaken off. MALDI mass spectra were recorded in a time-lag focusing linear time-of-flight mass spectrometer with a 1-m flight tube. Mass spectra shown in this work represent the sum of 20 to 100 individual spectra. Spectra were calibrated externally with the use of bradykinin, bovine insulin chain B, and equine cytochrome C as the calibrants.

ERDEC Experiment. Bacterial extracts of E. coli (ATCC 11776) were prepared using a modification of the procedure reported by Krishnamurthy et al [4]. Dry, lyophilized bacteria (2-3mg) were placed in a microcentrifuge tube to which was added 75 µl each of the following solutions: 10 mM Tris buffer (pH 8.0), 1% SDS and 0.1 mM β-mercaptoethanol. This mixture was vortexed then incubated at 95 °C for 20 minutes. The cloudy solutions were cooled to room temperature, 10 µl DNAse(I) (1mg/ml) was added and the solutions were held at room temperature for 20 minutes. The mixture was centrifuged at low speed for 10 minutes and the clear supernatant was transferred to a clean microcentrifuge tube. The soluble proteins were precipitated by addition of 1 ml of cold methanol followed by storage at 0 °C for 20 minutes. The proteins were pelleted by centrifugation at high speed for 10 minutes. The supernatant was removed and the precipitate was air dried for 10 minutes. Eight samples were reconstituted with 75 µl Tris (1mM, pH 8.0), combined and dried using a Centrivap.

For MALDI analysis the dried precipitate was reconstituted in 100 µl water. The resulting solution was centrifuged at low speed to remove any particulates before use. The MALDI samples were prepared by adding 1 µl of the extract solution to 9 µl of matrix solution (saturated α-cyano-4-hydroxycinnamic acid solution in 50:50 acetonitrile:0.3% TFA). This solution was vortexed for 30 seconds and 1 µl was immediately spotted onto a clean MALDI sample pin and allowed to air dry. Spectra were recorded on a Vestec MALDI-TOF in linear mode. The spectra were externally calibrated using a matrix ion peak (dimer, m/z 379.2) and cytochrome C (m/z 12361.5).

3. Results and Discussion

Figure 1 shows the MALDI mass spectra of E. coli (ATCC 9637) obtained using different solvents for matrix/sample preparation. In both spectra, the cell suspension was prepared in 0.1%

TFA and the first matrix layer on the probe tip was prepared using HCCA. The second-layer solution used for generating Figure 1A consisted of saturated HCCA in 33% acetonitrile/67% water and the sample solution (1:1) (all by volume). For Figure 1B, the second-layer solution consisted of saturated HCCA in 17% formic acid/33% isopropanol/50% water (FPW) and the sample solution (1:1) (all by volume). The common peaks observed in both spectra are labeled with asterisks. Figure 1 illustrates that the relative intensities of these peaks are different for the two spectra. Most peaks detected in Figure 1A appear in Figure 1B as well. Similar results were obtained from the other bacteria samples studied. The salient feature is that the use of FPW for preparing the sample provides more sensitive detection. A number of common peaks, albeit with different relative intensities, are observed under different sample/matrix preparation conditions. It appears that, with the two-layer method, variation of the solvent conditions for matrix/sample preparation has a much greater effect on the relative intensities of the peaks than on the actual number of proteins detected.

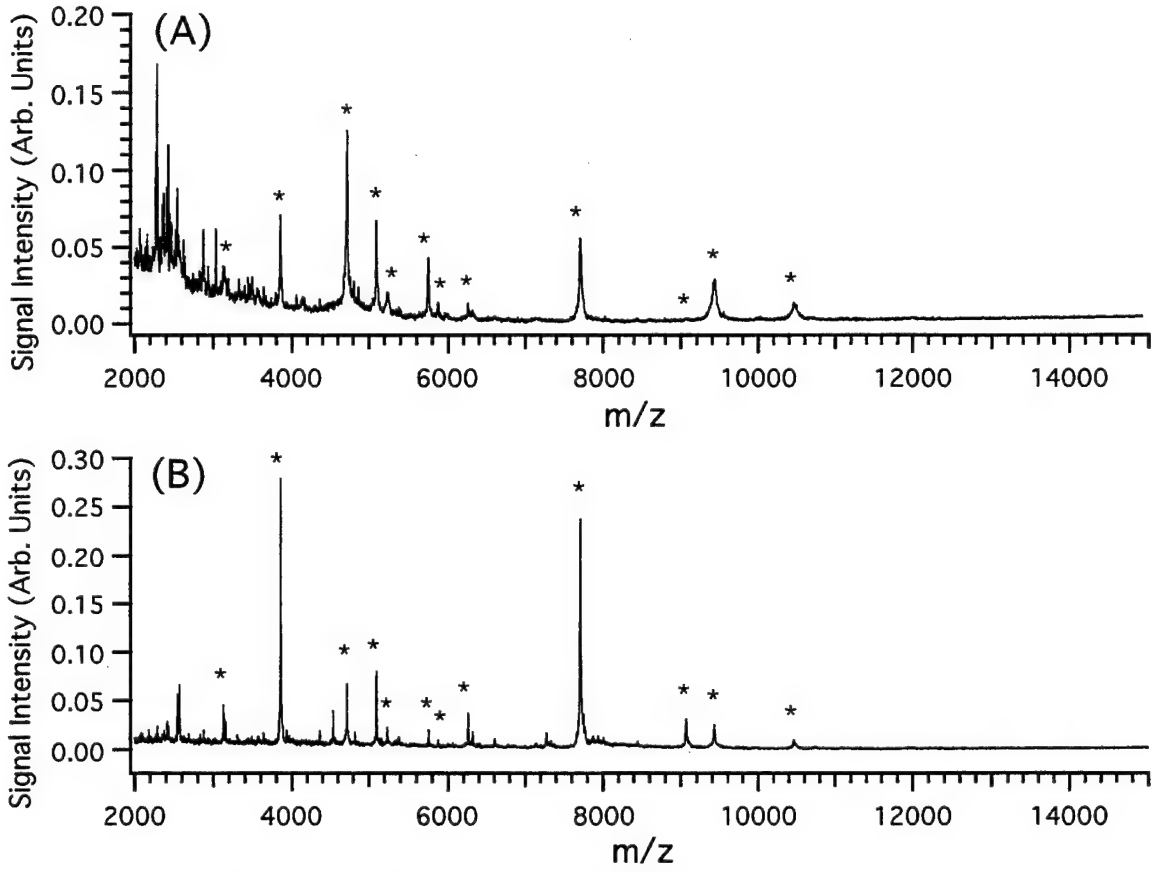


Figure 1. MALDI spectra of E. coli (ATCC 9637) obtained by using different sample/matrix preparation conditions.

The observation of differences in overall detection sensitivity and relative peak intensity under different sample preparation conditions is not surprising. The analyte incorporation and distribution in the matrix crystals can be affected by sample preparation [5]. For example, the variation of pH of the sample and matrix solution alone can significantly affect MALDI spectral patterns of protein mixtures [6]. Since bacteria identification in the MALDI method is based on one or more characteristic peaks in the spectrum, once the biomarkers are identified for a particular bacteria, optimal sample/matrix preparation methods can be designed for the sensitive detection of these biomarkers. While a universal sample preparation method for detecting all biomarkers for many different bacterial strains and/or species is desirable, such a preparation protocol may be difficult to find. The utility of multiple sample preparation protocols with each optimized for a small number or a group of biomarkers may be a sensible approach. Multiple sample handling is a standard feature in commercial MALDI instruments. Automation of the sample preparation step is possible. With this in mind, we can again examine the MALDI spectra of E. coli shown in Figure 1.

The peaks shown around m/z 7700 and 3800 are the dominant peaks in Figure 1B. If these peaks are to be used for identification of E. coli, the experimental conditions used for generating Figure 1B are preferred to those used for Figure 1A, from the MALDI detection point of view.

For the different bacteria samples examined, MALDI spectra obtained by using the two-layer method are quite different. For example, Figure 2 shows the spectra of bacillus cereus (ATCC 13824) obtained under the same conditions as those used for Figure 1. The common peaks are labeled with asterisks. Again, relative intensities of these peaks are different for Figure 2A and 2B obtained with the use of different solvents for sample/matrix preparation. The mass spectrum shown in Figure 2B displays two prominent regions of peaks at m/z around 7300 and m/z 3700. These peaks are different from those shown in Figure 1B. They may potentially be used as the biomarkers for bacillus cereus, if these peaks are found to be independent of other environmental changes such as growth culture media, etc.

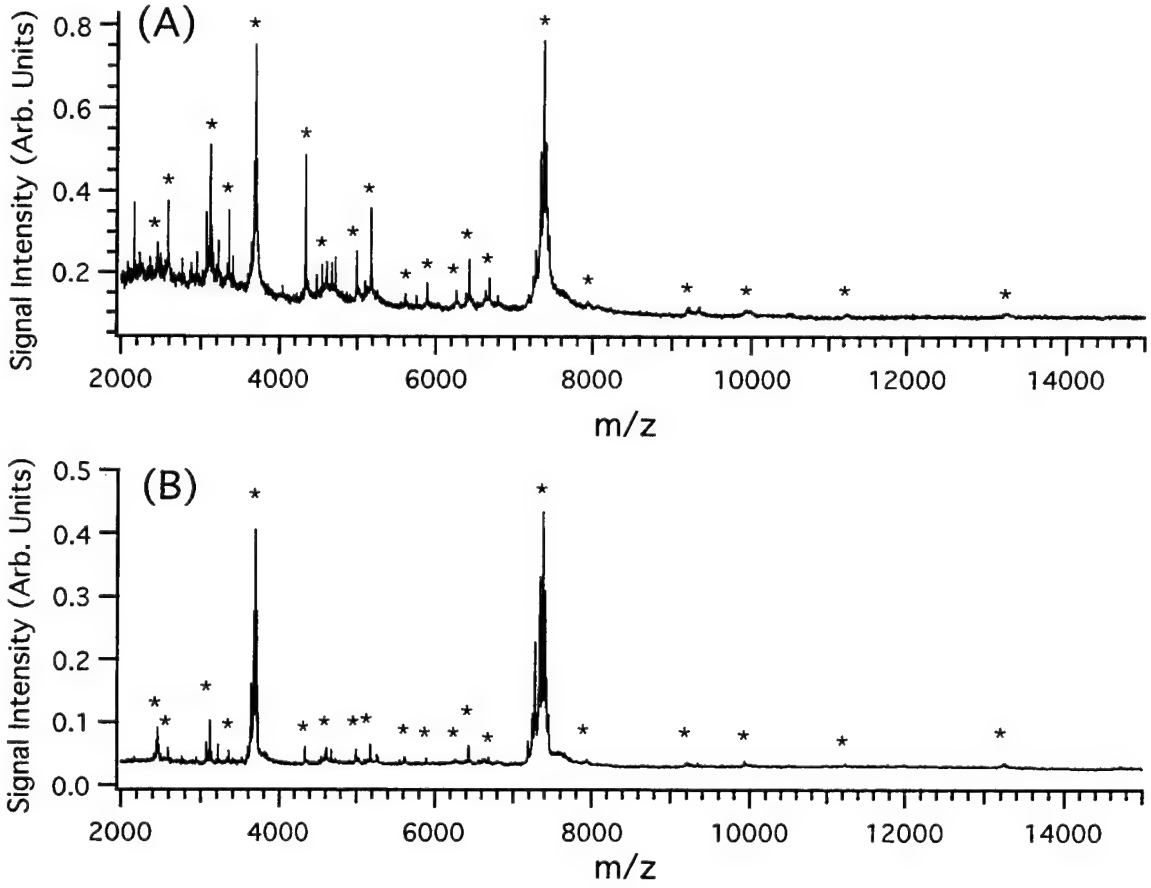


Figure 2. MALDI spectra of bacillus cereus (ATCC 13824) obtained by using different sample/matrix preparation conditions.

Using the solvent suspension method for protein extraction, it is found that mass spectral patterns can be significantly affected by the type of the suspension solvent used. This may be due to the solubility difference of the proteins in different solvents. Figure 3 shows the mass spectra of E. coli (ATCC 9637) obtained with two different extraction solvents, but under the same matrix/sample preparation condition. Figure 3A was obtained by using 0.1% TFA as the suspension solvent for protein extraction. The second-layer solution consists of 17% formic acid/33% methanol/50% water (FPW) and the sample solution in 0.1% TFA (1:1) (all by volume). For Figure 3B, a mixture of 17% formic acid, 33% methanol, and 50% water (by volume) was used for extraction. The second-layer solution was prepared with its solvent composition the same as that used for Figure 3A.

Figure 3 clearly shows that the extraction solvent can have a significant effect on the spectral patterns. There are a number of common peaks present in both spectra (labeled with arterisks). They are likely from the same proteins dissolved in both solvents. Comparing these two spectra also reveals that different proteins are extracted from the two solvent systems. This example illustrates that the selection of the extraction conditions is important for detecting the potential biomarkers. We also observed different patterns from other extraction solvent systems such as NH_4HCO_3 solution. A detailed account of these studies will be presented elsewhere.

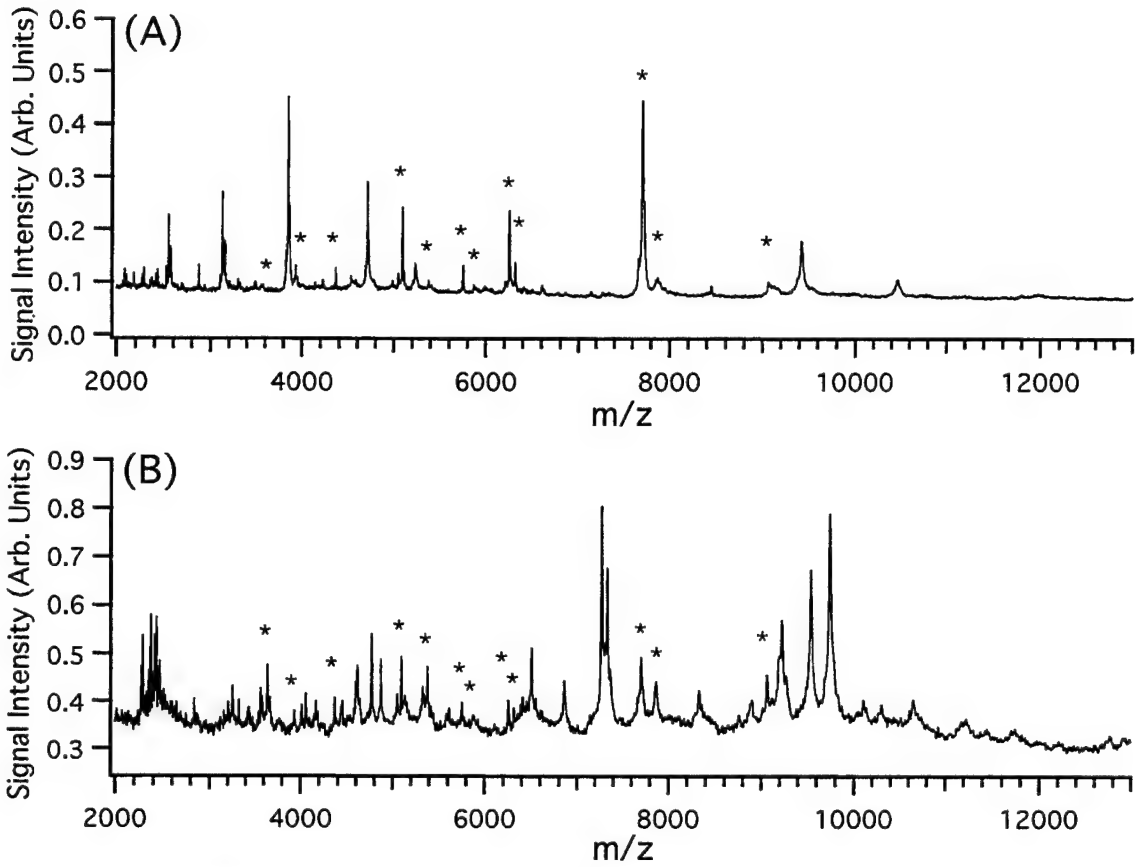


Figure 3. MALDI spectra of E. coli (ATCC 9637) obtained by using the same extraction methd, but different solvents.

Another method of protein extraction [4] was investigated using the same bacteria as those used in the solvent suspension method. Figure 4 shows the mass spectra of E. coli (ATCC 11776) obtained by the two extraction methods. Figure 4A was obtained using the same experimental condition as that used in Figure 1B for E. coli (ATCC 9637). Note that the mass spectral pattern shown in Figure 4A is quite different from that of Figure 1B. Many more peaks are detected in Figure 4A than in Figure 1B. Some of these peaks may be potentially used as the biomarkers for differentiating these different strains. Figure 4B shows the mass spectrum obtained using the procedures described in the Experimental section (i.e., ERDEC Experiment). Most peaks detected in Figure 4B are also found in Figure 4A, despite the fact that completely different sample preparation procedures were followed and different MALDI instruments were used to obtain the two spectra. It is interesting to note that the spectrum shown in Figure 4B resembles that shown in Figure 3B, although two different strains were used. It appears that only subtle differences in both number and content of proteins with masses ranging from 2000 to 13000 exist in these two strains of E. coli cells. The use of a proper extraction method can reveal such a difference, as shown by comparing Figure 3A with Figure 4A. These results underscore the importance of using a set of well-controlled experimental conditions for spectral recording and comparison. This study also

demonstrates that similar types of proteins can be detected from the same sample from different laboratories.

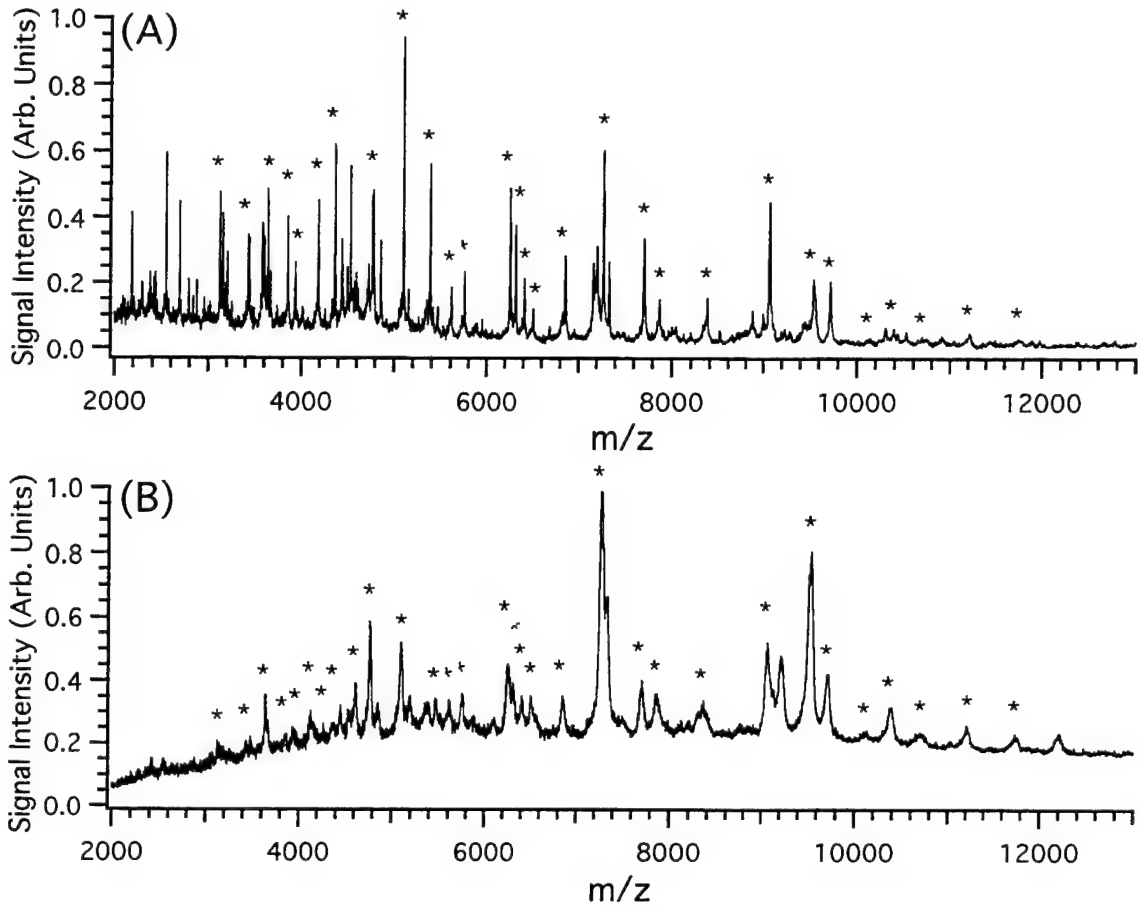


Figure 4. MALDI spectra of E. coli (ATCC 11776) obtained by using different extraction methds.

4. Conclusions

Matrix/sample preparation and the protein extraction process can have significant effects on the mass spectral patterns. The ability to use different solvent systems for extracting different proteins provides selectivity. This selectivity can potentially be useful in biomarker selection as well as in developing optimal sample preparation protocols for biomarker detection. This selectivity also dictates that a set of well-controlled experimental conditions need to be followed for spectral comparison. However, many common peaks from the same sample can be detected from different sample preparation procedures or from different laboratories.

5. References

1. P.O'Farrell, *J. Biol. Chem.* 250, 4007 (1975).
2. T.C. Cain, D.M. Lubman and W.J. Weber, Jr., *Rapid Commun. Mass Spectrom.* 8, 1026 (1994).
3. R.D. Holland, J.G. Wilkes, F. Rafii, J.B. Sutherland, C.C. Persons, K.J. Voorhees, and J.O. Lay, Jr., *Rapid Commun. Mass Spectrom.* 10, 1227 (1996).
4. T. Krishnamurthy, P.L. Ross, and U. Rajamani, *Rapid Commun. Mass Spectrom.* 10, 883 (1996).
5. Y.Q. Dai, R.M. Whittal and L. Li, *Anal. Chem.* 68, 2494 (1996).
6. R.M. Whittal, L. Li, unpublished.

ENHANCED PERFORMANCE OF CONTINUOUS FLOW MALDI USING TIME-LAG-FOCUSING

D. Shannon Cornett and John Wronka
Bruker Daltonics
19 Fortune Dr.
Billerica, MA 01821 USA

ABSTRACT

The use of liquid matrices with matrix assisted laser desorption/ionization have traditionally been associated with low mass spectrometry performance, namely low resolution and sensitivity. The recent application of time-lag-focusing to MALDI-TOF has provided substantial improvement in performance when combined with the use of solid matrices. We combine TLF with a CF-MALDI interface which uses liquid matrices to gain a similar improvement in mass resolution on two instruments of different flight path and vacuum pumping capacity.

INTRODUCTION

Development of matrix assisted laser desorption/ionization (MALDI) as an online detection/identification strategy for biological agents would be greatly facilitated by the use of liquid-based matrices, particularly when incorporated as part of a continuous-flow (CF) sample introduction interface. Of the two most common techniques used for analyzing large biological compounds, MALDI and electrospray, MALDI is advantageous for examining complex samples in that it is more amenable to mixture analysis and is less sensitive to salts and other contaminates. On the other hand, preparation of a typical solid MALDI sample can require a number of procedures which are usually carried out by a devoted operator. These involve mixing of matrix and analyte solutions, deposition of this mixture onto a target and insertion of the target into the instrument. A number of manufacturers have introduced robotic sample preparation devices that will automatically mix samples and deposit an aliquot onto a MALDI target, thereby minimizing human intervention. However the majority of these devices are simply automatic pipetters which still require some actions to be performed by an operator e.g., to place tubes into trays and to program the desired target position. An idealized approach is to have a means of introducing analyte/matrix samples directly into the MALDI source which requires no human intervention except perhaps for the initial experimental setup and which is capable of sustained operation.

One means of accomplishing this task is to use a continuous flow interface for direct infusion of liquid matrix into the MALDI source. CF interfaces have long been discussed as a means of incorporating the advantages of MALDI into an on-line analysis strategy but, their development has been hampered because liquid matrices typically yield lower mass spectrometry performance than the more commonly used solid matrices. The recent application of time-lag-focusing to MALDI has resulted in a substantial improvement in mass spectrometry performance. For example mass resolution and measurement accuracy have increased by an order of magnitude. In the work described here we show that the performance enhancements afforded by TLF are retained when the MALDI source is fitted with a CF interface. A secondary goal of this work is to show that the CF interface can be accommodated by a small mass spectrometer system with a limited vacuum pumping capacity.

CONTINUOUS FLOW INTERFACE

In meeting these goals the first obstacle is the choice of liquid matrix. It must have low volatility, absorb laser light, facilitate ionization and solubilize the analyte. Generally the liquid matrices examined to date with MALDI were established for fast atom bombardment studies, glycerol and nitrobenzyl alcohol for example. Unfortunately both of the current instruments use nitrogen lasers which emit 337 nm light, a wavelength at which few liquid matrices absorb strongly. In an earlier collaborative effort with the University of Georgia, Bruker developed a prototype CF interface which uses a combination of alpha-cyano hydroxycinnamic acid (ACHA) and glycerol (or ethylene glycol) as light absorber and low-volatility matrix, respectively. The interface, shown in Figure 1, continuously infuses a 3-8% aqueous glycerol solution into the mass spectrometer. The analyte, present in a saturated solution of ACHA in the carrier solvent is injected into the carrier solvent prior to entering the vacuum.

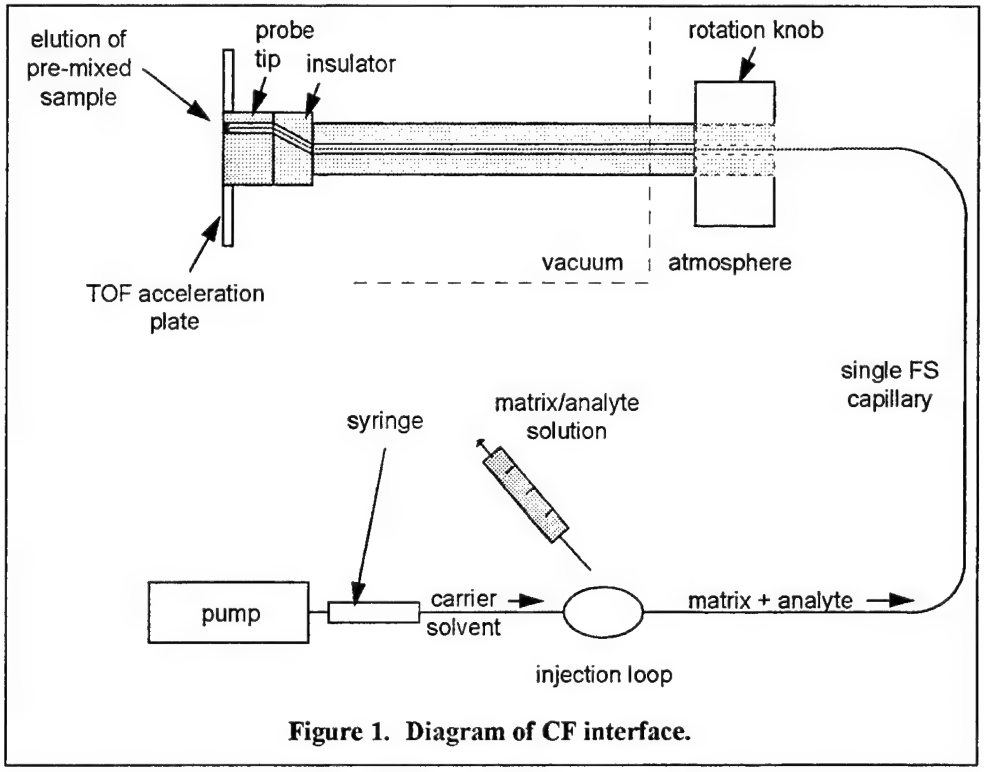


Figure 1. Diagram of CF interface.

In the course of these experiments the interface was fitted onto a 3m reflectron, (large) dual-pumped instrument and a 0.65m linear (small) singly-pumped instrument. Of primary concern was the level of pumping necessary to provide stable operational conditions for the high-voltage pulsing of the TLF source under the high gas load introduced by the CF interface.

LARGE SYSTEM

During development of the prototype CF interface the large instrument was not equipped for TLF measurements. However these early studies serve to provide a baseline level of instrument performance for observing the degree of enhancement provided by TLF. Without TLF this instrument previously demonstrated CF-MALDI mass resolution of 300-500 for analytes up to ca. 20 kDa when operated in reflectron mode. Linear resolution provided by CF operation was considerably lower, due in large part to a customized linear detector which optimizes signal intensity at the expense of mass resolution. With the addition of TLF to this instrument CF-MALDI mass resolution is increased 5-fold to 2,000-2,500 for m/z 1,000-2,000 in reflectron mode. Figure 2 is a reflectron CF-MALDI mass spectrum obtained from a mixture of the peptide standards, bradykinin, substance p and bombesin. A 3µl aliquot of sample (15 pmol/µl) was injected at a CF flow rate of 10 µl/hr.

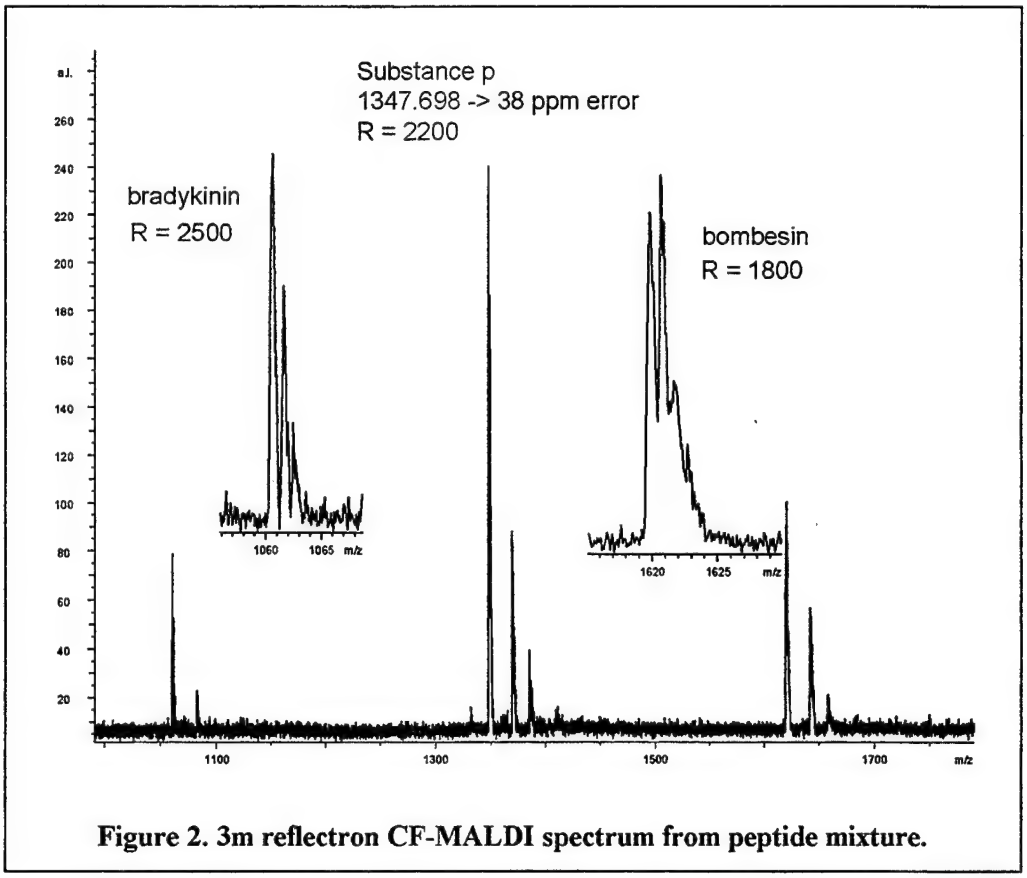


Figure 2. 3m reflectron CF-MALDI spectrum from peptide mixture.

In order to maintain as direct a comparison as possible between the large and small instruments, both were operated at 10 kV total acceleration voltage. This voltage provides an acceptable compromise between sensitivity and minimal arcing due to the higher background pressure produced by the infusing solvents. Before installation of the CF interface onto the large system the optimum operating conditions for 10 kV TLF were established using the standard MALDI probe. A mixture of peptides, bradykinin, angiotensin I, substance P and bombesin, were prepared as a solid MALDI sample using the HCCA matrix, and ion source voltages were empirically adjusted to produce optimum mass resolution across the mass range of the component ions. Under these conditions the *STANDARD* MALDI sample produced mass resolution of about 2,000. So from Figure 2 it is apparent that the CF interface is capable of achieving mass resolution similar to that achieved with solid MALDI samples. Admittedly the sensitivity demonstrated in Figure 2 is quite high compared to the levels achieved with solid samples, but the high signal to noise indicates that the CF spectrum in the figure is not sample limited.

The CF performance for higher molecular weight compounds was also examined in reflector mode. Figure 5 shows a reflectron CF spectrum from a 20 pmol injection of insulin. Although the mass resolution calculated for the insulin molecular ion is 600 it is important to note that without the capability of TLF it would be very difficult to resolve the ion formed by loss of water $[M-17]^+$. This data clearly demonstrates that the addition of TLF to this instrument provides substantial improvement in resolution for CF mass spectra.

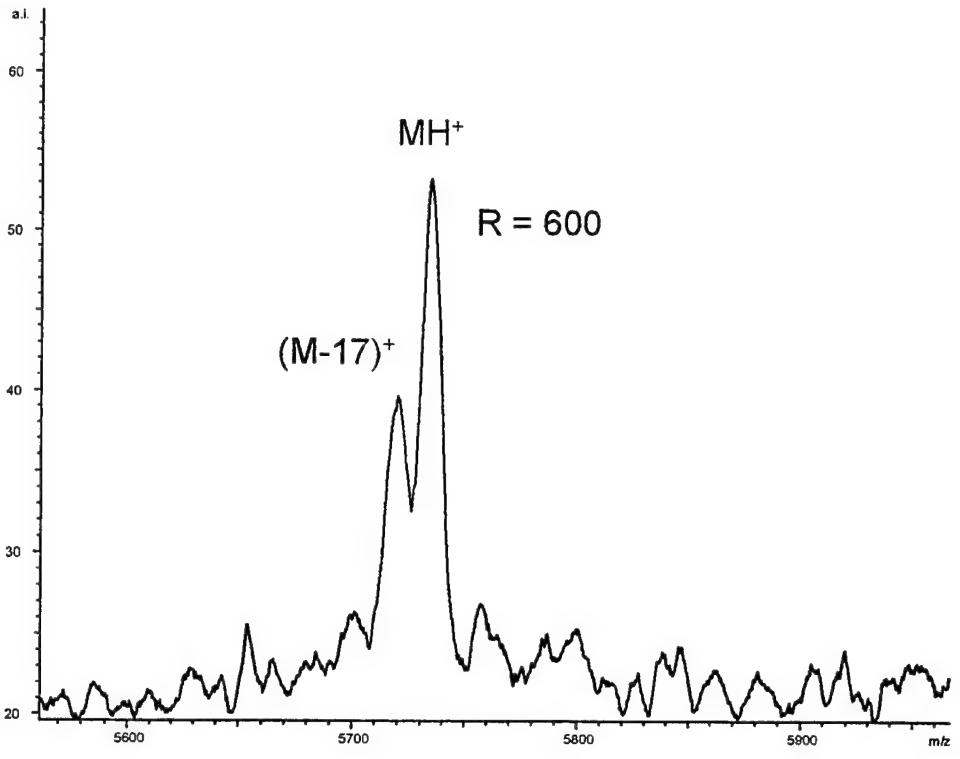


Figure 3. 3m reflectron CF-MALDI spectrum of insulin.

Before concluding experiments on the large system, a lysed sample of *E. Coli* was analyzed using the CF interface. The concentration of the sample was unknown but similar samples have been previously analyzed by solid MALDI with good results. Figure 4 shows a linear CF mass spectrum obtained from a 5 µl injection of this sample prepared with HCCA matrix. Although the signal/noise is not as high as that obtained for the protein standards, it should be noted that it was prepared from biologically buffered solution. The high concentrations of salts could contribute significantly to the high background and lower molecular ion intensity.

SMALL SYSTEM

Dramatic fluctuations in pressure due to sputtering was the major difficulty in working with the smaller system. Since these experiments were completed first we were not able to take advantage of using ethylene glycol as a solvent component. As such all data acquired on the small instrument was done using glycerol-water as the CF carrier solvent. Early experiments were made using varying ratios of glycerol in an attempt to gain some degree of control over the evaporation rate of the solvent as it eluted from the capillary in vacuum. Ratios from 3-10% were examined with 4-5% being optimum. Lower concentrations of glycerol offered little protection from solvent freezing due to evaporative cooling while higher concentrations increased the solution viscosity such that it was difficult to maintain reproducible flow rates through the 50 µm capillary. Even with a concentration of 4% glycerol maintaining a stable flow through the open-end capillary of the CF interface was difficult.

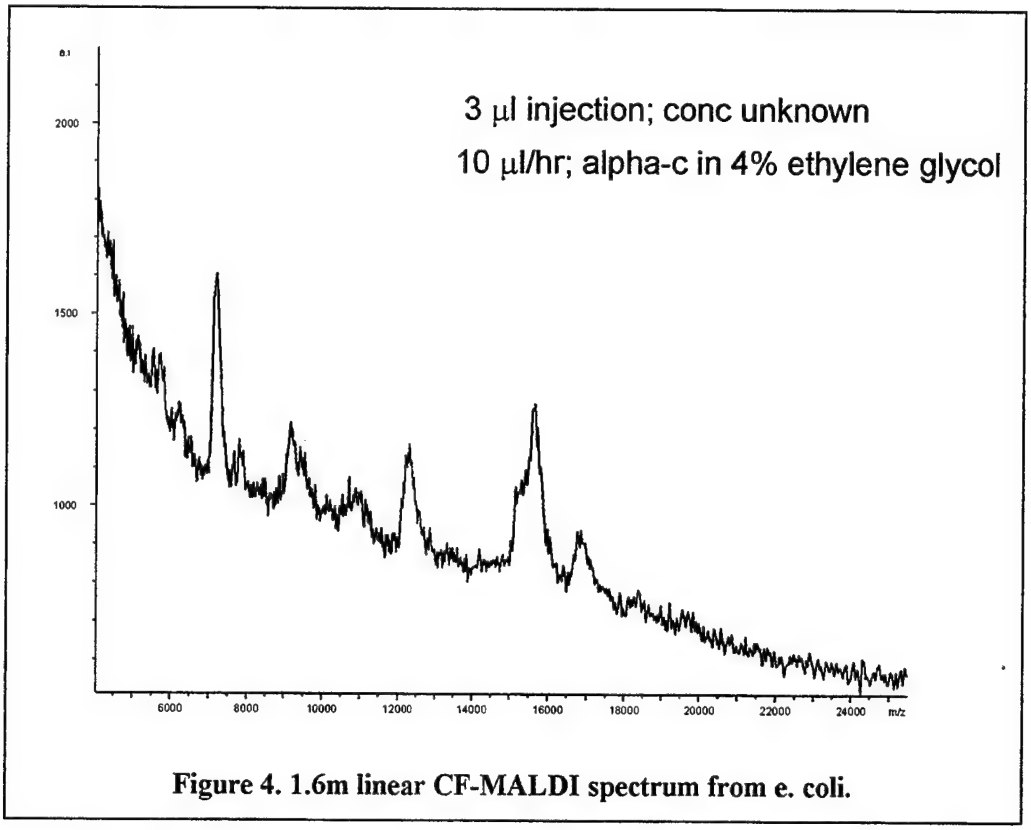


Figure 4. 1.6m linear CF-MALDI spectrum from e. coli.

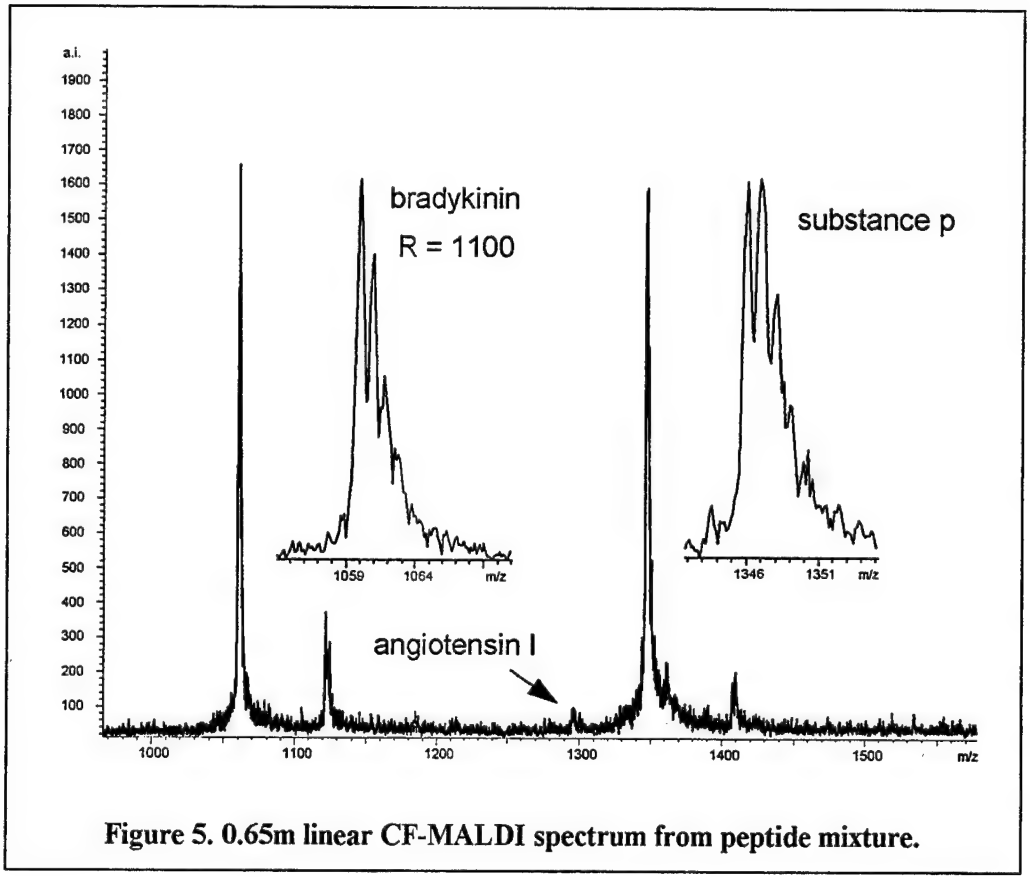


Figure 5. 0.65m linear CF-MALDI spectrum from peptide mixture.

Modifying the interface so that a short piece of braided nylon thread was positioned against the capillary opening of the interface provided a more stable vacuum inside the small system. Figure 5 was acquired from a

mixture of bradykinin, substance P and bombesin. The sample solution was prepared with each component present at a concentration of 10 pmol/μl with 3 μl being injected at a flow rate of 45 μl/hr. A total of 30 pmol loading represents a factor of 3-5 better sensitivity than previously obtained for CF-MALDI on this instrument. Additionally, for the two major peaks, bradykinin and substance P (inset views), the mass resolution is sufficient to partially resolve the ^{13}C isotopes (resolution ~1500).

CONCLUSIONS

These experiments show that time-lag-focusing can be used to improve the mass resolution achieved using a continuous-flow MALDI interface. Under similar instrumental conditions CF-TLF-MALDI is shown to produce mass resolution comparable to that obtained using solid MALDI samples. It is also shown that the operation of the CF interface is not limited to 'high end' instruments. The vacuum requirements of both the TLF ion source and CF interface can be met using a 'bench top' instrument. These results show promise for a simple, on-line CF-MALDI instrument.

ACKNOWLEDGEMENTS

The authors acknowledge the collaborative efforts of the University of Georgia and financial support from the U.S. Army.

ELECTROSPRAY/ION TRAP MASS SPECTROMETRY FOR THE DETECTION AND IDENTIFICATION OF ORGANISMS

Scott A. McLuckey and James L. Stephenson, Jr.
Chemical and Analytical Sciences Division
Oak Ridge National Laboratory
Oak Ridge, TN 37831-6365

ABSTRACT

Current electrospray ion trap methodology for rapid mixture analysis of proteins used for the identification of microorganisms is described. Development of ion/ion reaction techniques (e.g. reactions of multiply-charged protein cations with singly-charged anions) from both a fundamental and practical approach are presented, detailing the necessary steps and considerations involved in complex mixture analysis. Data describing the reduction of the initial charge states of electrospray ions to arbitrarily low values, the utility of ion/ion reactions for mixture separation on the millisecond time scale, and effects of excess singly-charged reactants on detection and storage efficiency are illustrated.

INTRODUCTION

Recently, the combination of electrospray (ES) ionization with ion trap mass spectrometry, has proven to be a valuable tool in solving bioanalytical problems (1-4). The extension of ES mass spectrometry to the field of microbiology (e.g. viral analysis) for the identification of organisms is now becoming a reality (5-6). Despeyroux and coworkers have shown that the protein content of cricket paralysis virus can be mapped using intact virus without a preliminary extraction or sample preparation step (5). In addition, the multiple charging phenomenon of ES has made possible the analysis of whole tobacco mosaic virus and rice yellow mottle virus particles which have molecular weights in excess of 40 million Da (6). This multiple charging phenomenon associated with ES yields a distribution of charge states, thereby allowing for multiple mass-to-charge measurements for a single species, and effectively reducing mass-to-charge values to a range amenable to most mass analyzers. Although the multiple charging phenomenon facilitates biopolymer mass measurement in some respects, complications can easily arise when ions derived from more than one biopolymer are present. Since multiple charging compresses the "mass scale", the distance between adjacent charge states on the m/z scale decreases with increasing charge. Therefore this mass scale compression can make difficult the resolution of mixture components with modest differences in mass-to-charge ratio (7).

Currently, two strategies exist to address peak overlap in ES mass spectra. The first centers on increasing the resolving power of the mass analyzer, utilizing Fourier transform ion cyclotron resonance (FTICR) mass spectrometry. The FTICR permits far more detail to be extracted from mixtures of ions derived from ES (8) than from the more commonly used quadrupole based techniques. Secondly, computer algorithms have been developed to extract mass information directly from ES mixture data (9), however the quality of the data is still dependent on the ability of the mass analyzer to resolve adjacent charge states (7).

We have developed a novel ion/ion reaction strategy to decompress the "mass scale" by reducing the charges on the ions to the point at which the mass analyzer can resolve the mixture components. In these experiments, multiply charged protein cations are reacted on the millisecond

time scale, with singly-charged anions in order to reduce the charge states (e.g. increase the distance between adjacent peaks) of the protein cation population via proton transfer to arbitrarily low values (7,10-13). The extent to which the mass scale can be decompressed is limited by the upper limit to the mass-to-charge scale of the ion trap mass spectrometer. Some advantages to this approach for mixture analysis include well-defined ion/ion chemistry (exclusive proton transfer reaction products observed), minimal ion/neutral effects (minimal scattering, ion/molecule reactions), and fast reactions which maximize speed/duty cycle. This ion/ion reaction phenomenology will be employed in conjunction with targeted protein analysis, for the identification of microorganisms.

EXPERIMENTAL

All experiments were carried out with a home-made electrospray source coupled with a Finnigan MAT ion trap mass spectrometer, modified for injection of ions formed external to the ion trap through an end-cap electrode. Details of the electrospray /ion trap interface have been described (14). Further modifications have been made to this apparatus to allow anions to be injected through a 3 mm hole drilled through the ring electrode. Details of these modifications and others made to facilitate analysis of high mass-to-charge ratio ions are reported elsewhere (12). A brief description of the hardware changes is given here.

An atmospheric sampling glow discharge ion source has been mounted on a side port of a 6 in. cube used to support the ion trap assembly. The ion trap is situated such that there is a line of sight from the exit aperture of the glow discharge ion source to the 3 mm hole in the ring electrode. A lens stack is mounted off the glow discharge ion source to facilitate ion transport to the ring electrode. The discharge is pulsed under software control. The output of the pulser is connected to the anode of the glow discharge source. The pulser acts as a fast switch which alternates between a voltage sufficient to strike a discharge and ground. This arrangement allows for independent control of cation and anion accumulation. The glow discharge provides a convenient means for forming a wide variety of singly charged anions in high abundance. In all cases, the multiply charged protein ions were accumulated prior to anion accumulation.

Cations were injected axially into the ion trap for periods ranging from 0.02 to 0.2 s. In all cases, helium was admitted into the vacuum system to a total pressure of 1 mtorr. Anions were formed by sampling the headspace vapors of perfluoro-1,3-dimethylcyclohexane (PDCH) into the glow discharge operated at 800 mtorr. For all ion/ion reactions described, anion accumulation periods ranged from 10 to 30 ms. Mutual storage times for ion/ion reactions varied from 30 to 350 ms. Mass range extension of the ion trap as high as m/z 70,000 was accomplished using the resonance ejection technique.

Purified tobacco mosaic virus in a 0.1 M phosphate buffer solution was obtained from American Type Culture Collection. The virus was isolated using a 10 kDa molecular weight cut-off filter, and was washed with several 1 ml aliquots of HPLC grade water. The virus was then resuspended in a 50:50 water/methanol solution containing 1% acetic acid. This solution was found to significantly disrupt the noncovalent interactions of the coat protein with the viral RNA. The proteins bovine serum albumin, transferrin, and ubiquitin were obtained from Sigma Chemical Company. PDCH was purchased directly from Aldrich Chemical Company.

Figure 1 shows the electrospray mass spectrum of the tobacco mosaic virus coat protein in the absence of anions, and after ions have been subjected to reaction with anions derived from PDCH for 100 ms. Under these reaction conditions ion/ion proton transfer reactions can reduce the charge on the TMV coat protein to arbitrarily low values. Previous results in our laboratory have shown that ion/ion rate data exhibit a linear dependence of rate on the square of the protein charge. These results suggest that the ion/ion reactions proceed at constant efficiency for all charge states and can therefore be used to reduce the charge of multiply-charged protein ions to arbitrarily low values.

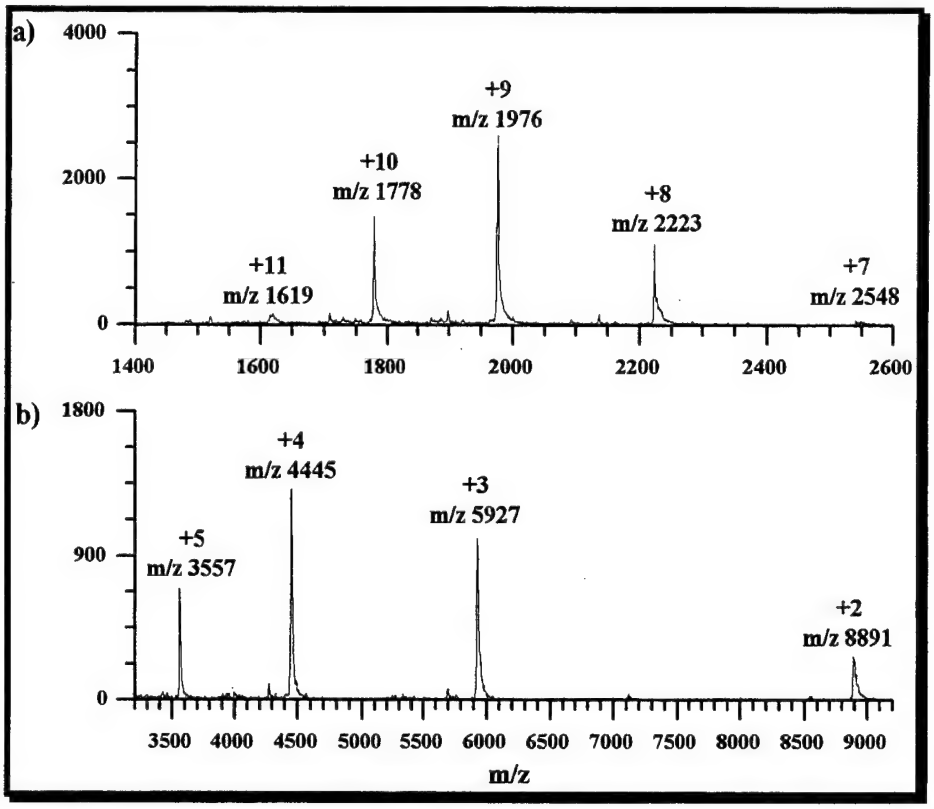


Figure 1 (a) ES mass spectrum of TMV viral coat protein. (b) Ion/ion reaction with PDCH anions for 100 ms.

In **Figure 2a** is shown the ES mass spectrum of a mixture containing ubiquitin, bovine serum albumin (BSA), and bovine transferrin. The latter two proteins were present in the mixture at approximate concentrations of 10 and 5 pmol/μL respectively. The ubiquitin component was present from contamination in the line leading to the electrospray needle and was used to provide an internal calibration of the mass-to-

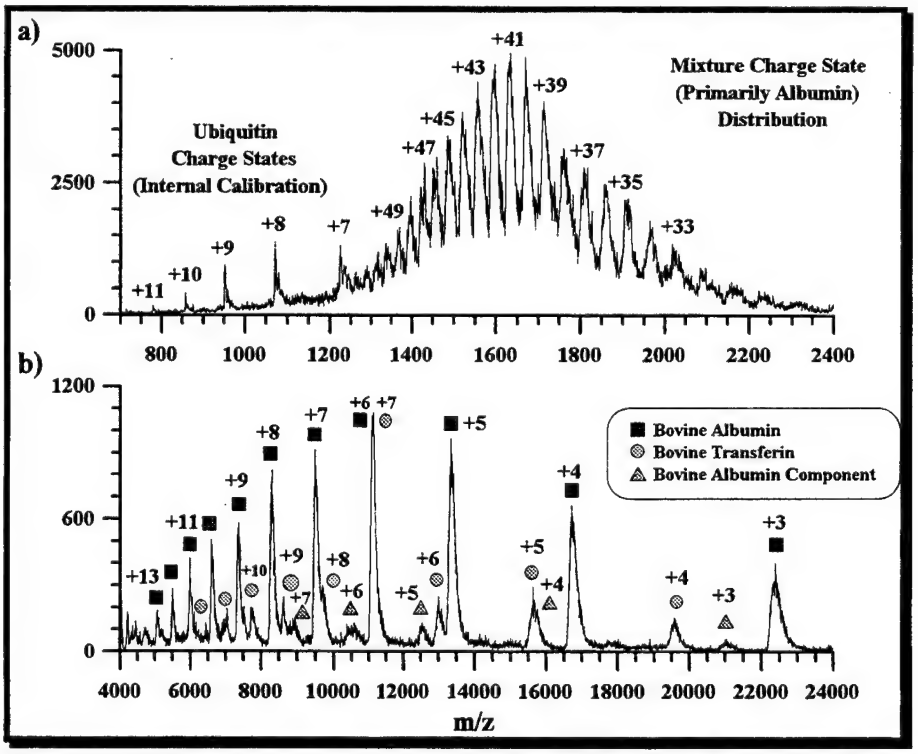


Figure 2 (a) ES mass spectrum of a protein mixture containing ubiquitin, bovine serum albumin, and bovine transferrin. (b) Ion/ion reaction of the mixture with PDCH anions for 295 ms showing resolution of mixture components.

charge scale. While the presence of BSA and ubiquitin is clear from casual inspection of **Figure 2a**, contributions to the spectrum from transferrin are not obvious. The total ion signal derived from equimolar solutions of each protein run separately shows that BSA gives a much greater response under identical ES conditions. The transferrin ions are buried within the BSA cation signal, probably for this reason. Figure 2b shows the spectrum that results from the reaction of the ions of Figure 2a with anions of PDCH for 295 ms. This example illustrates the mass scale decompression possible from charge state reduction and makes clear the presence of at least three distinct charge state distributions. One corresponds to BSA, one to transferrin, and one to an unknown mixture component derived from BSA.

While the chemistry associated with ion/ion reactions is clearly of primary interest in biological mixture analysis, the physics associated with the mutual storage of oppositely charged ions clouds can also be expected to affect important aspects of the ion trap experiment. For high number densities of singly-charged anions present during the ion/ion reaction experiment, two major phenomena affecting the physics of the stored ions associated with the mutual storage of cations and anions have been observed. First, the attractive potential experienced by the high mass cations arising from the presence of a large number of singly-charged anions can augment the trapping potential afforded by the quadrupole ion trap. **Figure 3a** shows the ion/ion reaction spectrum of ovalbumin (mw ~ 42,500 Da) after a 152 ms reaction time with anions derived from PDCH. Note that the primary peaks observed are the $[M+3H]^{3+}$ and the $[M+2H]^{2+}$ cations, with no response observed for the $[M+H]^{+}$ ion. In this example, the mass analysis conditions are set such that any ions with a m/z value greater than 3×10^4 will "leak" out of the analyzer due to insufficient pseudopotential well depths (e.g. little or no signal observed for the singly-charged ovalbumin).

In **Figure 3b** the same experiment is performed, except that the PDCH anions were present during the mass analysis period. Note that intense signals are observed for the anions (ejected via mass selective instability) despite the fact that the detector voltages are optimized for cations, and that the observed anion signal is roughly three time larger than the observed ovalbumin cation data (thus demonstrating the significant excess of anions). The $[M+H]^{+}$ ion of ovalbumin gives rise to the most intense cation signal in **Figure 3b** whereas there is essentially no evidence for this ion in

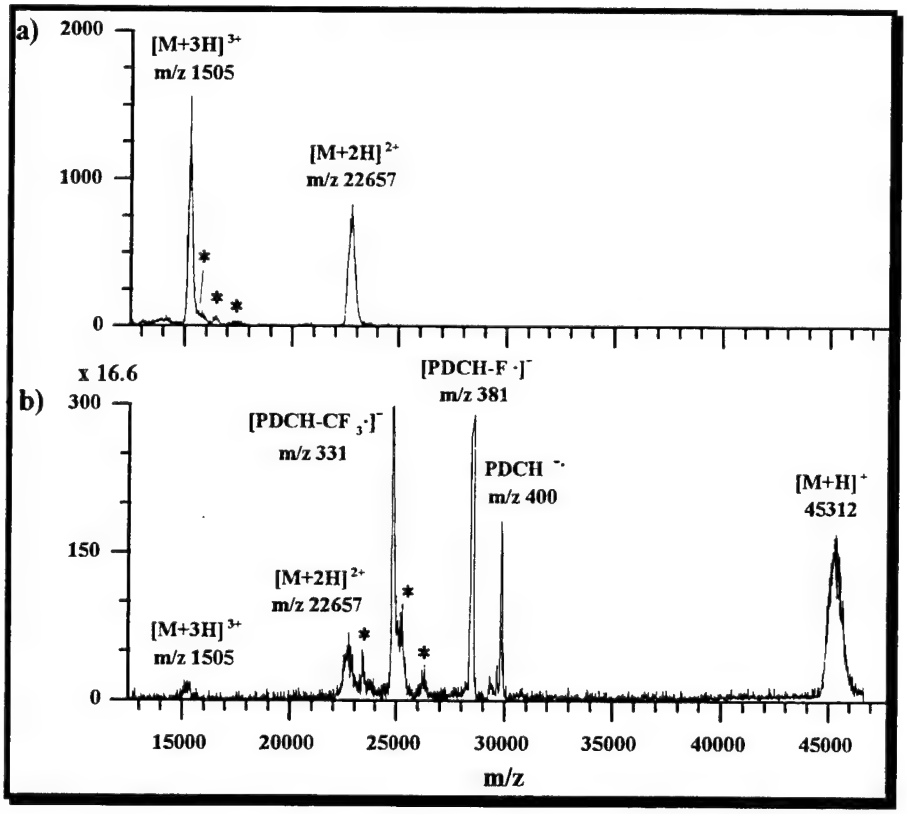


Figure 3 (a) ES mass spectrum of ovalbumin following 152 ms of mutual storage with anions derived from PDCH. The anions were then removed prior to mass analysis via resonance ejection. (b) ES mass spectrum acquired under the same conditions except that the anions were not ejected prior to mass analysis. Cations were ejected via resonance ejection whereas anions were ejected via mass selective instability. (*) Indicates peaks arising from unidentified impurities in the ovalbumin sample.

Figure 3a. For this example, high m/z cations can be trapped by the electric field created by an anion cloud when the trapping potential created by the oscillating quadrupole field is insufficient to do so. This creates the interesting situation in which the quadrupole field traps the anions which, in turn, are largely responsible for trapping the high mass cations. This phenomenon can then be exploited to store ions of higher m/z than might otherwise be trapped.

Another effect due to the presence of an ion cloud of opposite polarity and much higher total charge in the center of the trap can also be seen in **Figure 3**. For the ion/ion reaction spectrum in **Figure 3b**, signals arising from the +3 and +2 charge states are observed at less than 1% and 6% respectively of their intensities in **Figure 3a**. In this case the anion cloud which is present during the resonance ejection of the +3 and +2 charge states, interferes with the mass analysis of high m/z cations. For this reason it is desirable to eject the anions prior to resonance ejection of the cations (for the $[M+H]^+$ cation of ovalbumin, the anion cloud is effectively removed from the trap just prior to detection).

CONCLUSIONS AND FUTURE WORK

Ion/ion proton transfer reactions involving multiply-charged proteins can readily be effected in the quadrupole ion trap. These reactions can successfully reduce the charge states to arbitrarily low values, and separate mixtures of ions of different mass and charge but similar mass-to-charge ratios. This is a particularly common scenario in ES mass spectrometry of proteins and other biopolymers. Fluorocarbon anions, such as those derived from PDCH, appear to react exclusively via proton transfer with multiply protonated species, even for globular proteins as large as BSA and bovine transferrin. Such consistent and predictable behavior is important if useful mass information is to be derived from separating ion mixtures via charge state reduction.

For the reduction of charge states associated with ion/ion reactions of mixture components, ion storage and resonance ejection of the high m/z product ions can be significantly affected by the presence of an ion cloud of opposite polarity and much higher total charge in the center of the ion trap. The presence of the anion cloud can be used to help store high m/z protein cations which otherwise might not be trapped. In addition, the anion cloud can adversely affect the resonance ejection of mass analysis due to the fact that the electric field generated by the anions can rival or exceed the influence of the quadrupole field on cation motion.

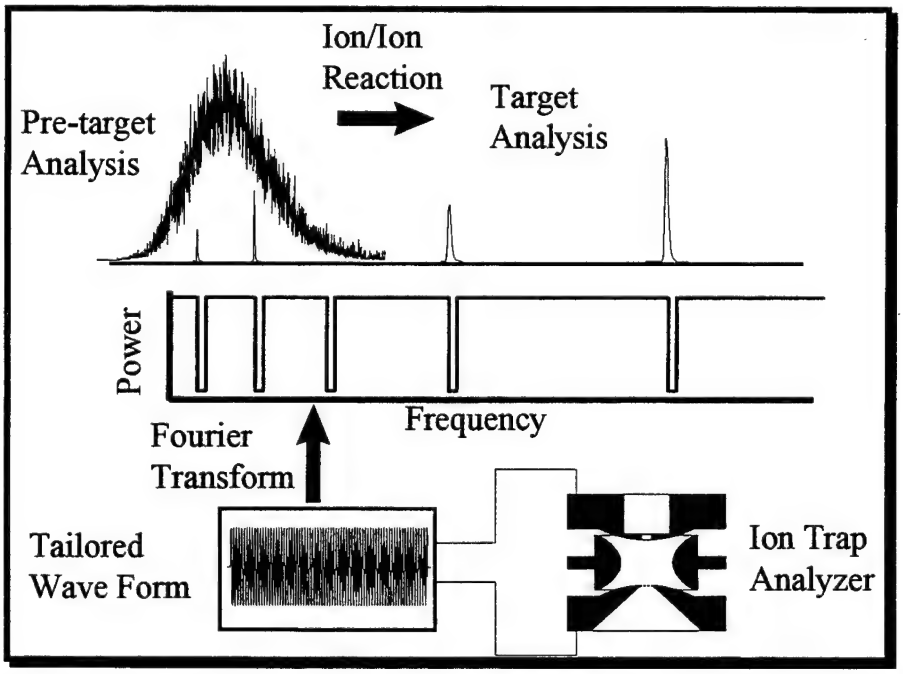


Figure 4 Experimental design for target analysis of marker proteins for species identification.

By understanding the underlying principles of the ion/ion reaction process, experiments can now be designed for the target analysis of marker proteins in various microorganisms for species identification. In **Figure 4**, the pre-target analysis represents the proteome of a given organism with the marker protein (unique to a particular genus species) shown "buried" under the starting protein mixture. In order to identify the microorganism, a tailored waveform is applied to the endcap electrodes which is specific for the charge state distribution of the marker peptide, ranging from the higher charge states generated by the ES ionization process to the lower charge states obtained via ion/ion reaction. This tailored waveform is set such that only ions of designated m/z values are stored while unwanted signal due to other proteins and matrix components are ejected from the trap. The next step in the experiment is to subject the ion signal present in the high charge state isolation windows to ion/ion proton transfer reactions with PDCH anions. After an approximate ion/ion reaction time of 100 ms or less, if peaks are present in the lower charge state windows, then an identification of the microorganism can be made.

ACKNOWLEDGEMENTS

This research was sponsored by the Office of Research and Development, U.S. Department of Energy, under contract DE-AC05-96OR22464 with Oak Ridge National Laboratory managed by Lockheed Martin Energy Research Corporation.

REFERENCES

(1) Ducret, A.; Bures, E.J.; Aebersold, R. *J. Prot. Chem.* **1997**, *16*, 323-328.
(2) Griffey, R.H.; Grieg, M.J.; Gaus, H.J.; Liu, K.; Monteith, D.; Winnman, M.; Cummins, L.L. *J. Mass Spectrom.* **1997**, *32*, 305-313.
(3) Wu, J.T.; Huang, P.Q.; Li, M.X.; Qian, M.G.; Lubman, D.M. *Anal. Chem.* **1997**, *69*, 320-326.
(4) Korner, R.; Wilm, M.; Morand, K.; Schubert, M.; Mann, M. *J. Am. Soc. Mass Spectrom.* **1996**, *7*, 150-156.
(5) Despeyroux, D.; Phillpotts, R.; Watts, P. *Rapid Comm. Mass Spectrom.* **1996**, *10*, 937-941.
(6) Siuzdak, G.; Bothner, B.; Yeager, M.; Brugidou, C.; Fauquet, C.M.; Hoey, K.; Chang, C.M. *Chemistry & Biology* **1996**, *3*, 45-48.
(7) Stephenson, Jr., J.L.; McLuckey, S.A. *Anal. Chem.* **1996**, *68*, 4026-4032.
(8) Winger, B.E.; Hofstadler, S.A.; Bruce, J.E.; Udseth, H.R.; Smith, R.D. *J. Am. Soc. Mass Spectrom.* **1993**, *4*, 190-192.
(9) Labowsky, M.; Whitehouse, C.; Fenn, J.B. *Rapid Comm. Mass Spectrom.* **1993**, *7*, 71-84.
(10) Stephenson, Jr., J.L.; McLuckey, S.A. *J. Am. Chem. Society*, **1996**, *118*, 7390-7397.
(11) Stephenson, Jr., J.L.; McLuckey, S.A. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 637-644.
(12) Stephenson, Jr., J.L.; McLuckey, S.A. *Int. J. Mass Spectrom. Ion Processes* **1997**, *162*, 89-106.
(13) Stephenson, Jr., J.L.; McLuckey, S.A. *Anal. Chem.* **1997**, in press.
(14) Van Berkel, G.J.; McLuckey, S.A.; Glish, G.L. *Anal. Chem.* **1990**, *62*, 1284-1295.

NOVEL MICROSPRAY METHODS FOR RAPID BIO-DETECTION BY MASS SPECTROMETRY

Terry D. Lee, Michael T. Davis, and Douglas C. Stahl
Beckman Research Institute of the City of Hope
Duarte, CA 91010, U.S.A.

Thaiya Krishnamurthy
U.S. Army, Edgewood RDEC
Aberdeen Proving Grounds, MD 21010, U.S.A.

ABSTRACT

A methodology is decribed for the rapid characterization of peptide and protein mixtures by on-line liquid chromatography - mass spectrometry (LC/MS). Key components of the technique are a micro-scale electrospray ionization source in which the LC column is integrated with the electrospray needle, a pressure controlled pumping system for solvent delivery, and an ion trap mass spectrometer. Each LC/MS analysis can be completed in 5 minutes with a total of ten minutes between successive sample injections. The technique was used for the identification of bacteria and gel separated proteins. A novel data processing technique which is capable of real time analysis of bacterial biomarker spectra is also described.

1. INTRODUCTION

The development of micro-scale electrospray ionization sources [1] [2] has greatly increased the sensitivity of mass spectral analyses of large biomolecules. At the same time, the technology of small scale, liquid chromatography separations has advanced sufficiently to be directly coupled to micro-electrospray (microspray) sources [3]. The result is a powerful analytical tool that can be used to characterize very complex biological mixtures with impressive sensitivity. Tandem mass spectrometry (MS/MS) methods have greatly increased both the quantity and quality of the information that can be obtained from a given molecular structure. Modern instrumentation developments (particularly ion trap analyzers) have greatly increased the ease with which MS/MS data can be collected, and techniques have been developed to efficiently perform such analyses during the course of an on-line LC separation [4].

One disadvantage of conventional LC/MS/MS approaches is the length of time needed to both collect and analyze the data. LC separations typically require 30-60 min to complete. This would seem to disqualify LC/MS approaches for applications where short intervals between sampling events is desirable, such as the detection and identification of biological threat agents. At the same time, the large quantity of data collected during an LC/MS run makes it difficult to extract the necessary information in a timely fashion. In this paper, we describe experimental approaches which greatly decrease the time needed to collect and analyze mass spectral data on complex mixtures of peptides and proteins. Using a novel gradient LC pumping system and microspray column, it is possible to achieve adequate separations of complex mixtures of peptides and proteins in 5-6 min, with a total elapsed time between sample injections of only 10 min. The analysis parameters of an ion trap mass analyzer can be adjusted to keep pace with the rapidly eluting components. A methodology for analyzing mass spectra is described that makes it possible to identify protein biomarkers as quickly as they are eluted from the LC column.

2. MATERIALS AND METHODS

2.1 HPLC and microelectrospray interface

All LC/MS analyses were performed using an Apple Macintosh controlled microcapillary HPLC system developed at the City of Hope and previously described by the authors [5]. The 150 µm ID x 350 µm OD on-line microspray needles used in these analyses were pulled using a commercial laser-based micropipette puller (Sutter Instrument Co. Novato, CA, model p-2000) to a terminal ID of approximately 5-10 microns. The final assembly of the microspray needle (Figure 1) was performed under a low power (6x) dissecting scope as previously described [3], except the second transfer line was withdrawn after the frit (Durapore) had been placed in its final position. The fritted tip was then coupled to a Valco 1/16" x 1/32" ZDV reducing union and packed at 4000 psi with Vydac, 5 µm, C_{18} packing. The packed tip was connected to a 75 µm ID x 350 µm OD transfer line using a PEEK capillary Tee (Valco) and graphite ferrules. A 0.3 mm Gold wire was introduced through the off-axis inlet to apply the electrospray potential. A 2-3 mm section of polyimide coating was removed from the 75 µm ID transfer line to form an on-column flowcell to monitor sample injections and the progression of the organic gradient.

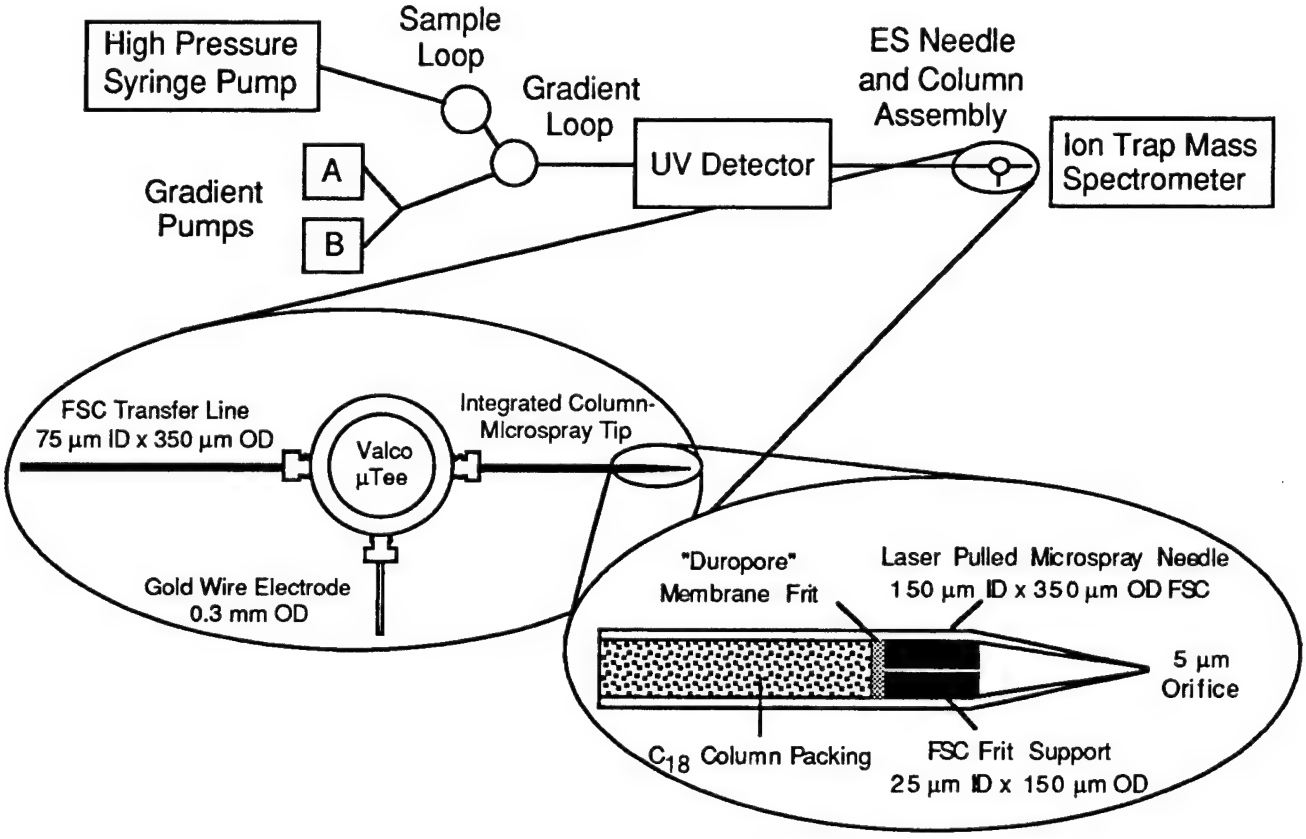


Figure 1. Apparatus for rapid LC/MS/MS analysis of Peptide and Protein Mixtures.

2.2 Mass Spectral Analyses

All mass spectra were acquired using a Finnigan MAT LCQ ion trap mass spectrometer. The microspray interface was mounted using a multi-axis translational positioner on a custom, dove-tailed, Plexiglas platform in place of the original Finnigan source as previously described [6]. For the analysis of the digested gel separated protein, the peptide CID spectra were searched against the OWL database [7] under enzyme constraints (trypsin) with a 1.5 amu parent and fragment mass tolerance. Data collected during the analyses of the bacteria samples were analyzed in real time using LabVIEW (National Instruments) based software running on the mass spectrometer data system computer.

3. RESULTS AND DISCUSSION

3.1 LC/MS System Components

The key components of the rapid LC/MS system (Figure 1) are a combination electrospray needle - chromatography column and an HPLC pumping system that has been optimized for low flow gradient deliveries. By incorporating the column within the electrospray needle, extra column dead volumes are eliminated and backpressures are reduced making it possible to achieve higher than normal flow rates. Typically, a 150 μm diameter column would be operated with a flow rate of approximately 1.0 μl/min. For the work described in this paper, a flow rate of about 1.3 - 1.5 μl/min was used. The small orifice of the electrospray needle (~5 μm) makes it possible to have stable ion emission over a broad range of flow rates (0.02 - 2.0 μl/min) and solvent compositions (0 - 100% organic) without the use of sheath liquids or sheath gas.

The LC separation was performed using a pumping system designed to deliver a preformed gradient without split flows. The pump is operated in a pressure controlled mode which provides for instantaneous changes in flow rate. Thus, sample loading and column cleanup can take place at flow rates (13-15 μl/min) as much as ten times that used for the separation. Because the uv detector is located before the column, it only provides information about the operation of the LC pumping system. By noting the arrival time of the gradient at the uv flow cell, it is possible to obtain a measure of the flow rate.

3.2 Ion Trap Parameters for Rapid Protein Analyses

For the analyses reported here, the entire chromatographic run is only 5 - 6 minutes with peak widths of only a few seconds. This imposes severe time constraints on the mass spectrometer and scan parameters were carefully chosen to optimize data collection. The rapid acquisition of full mass range spectra from 300 - 1850 m/z at a rate greater than 2 scans/second was achieved by reducing the micro scan count to one, disabling the automatic gain control (AGC), and using a fixed ion injection time of 50 ms. Under these conditions, the rapid, sensitive, analysis of a standard peptide digest mixture could be achieved using sample amounts ranging from 10 - 500 fmole (data not shown). The average peak width increased from 3.4 seconds (8 scans) at 10 femtomoles to 4.2 seconds (9.5 scans) at 500 fmoles. One consequence of using a constant ion injection time is the introduction of a significant shift (0.5 m/z) in the mass values for major components at the higher sample amounts.

More accurate mass values for individual peptide components could be obtained by using the zoom scan (narrow range, higher resolution scan) function of the LCQ. The instrument was programed to switch to the zoom scan mode whenever an ion with an intensity above a set threshold was encountered. The duration of the zoom scan analysis was extended to nearly 4 seconds by setting the microscan count to 20, which roughly corresponded to the mean peak width. The result of summing many zoom scans was a high quality spectrum of the isotope cluster and mass values accurate to within 0.15 amu (data not shown) for peptides from a standard cytochrome c Lys c digest mixture.

A similar approach was taken to obtain high quality MS/MS spectra. The instrument was programed to collect MS/MS spectra for each qualifying ion. The ion injection time was set at 500 ms, the charge state was set to a default value of 2, and the number of MS/MS scans for each event was set at five. The result of a rapid LC/MS/MS analysis of a one dimensional gel separated protein is given in Figure 2. The quality of the MS/MS spectra is very high even at sample levels below 100 fmole unless the ion is first detected after a majority of the peak has eluted.

3.3 Rapid Identification of Bacteria

A number of bacteria samples were analyzed using the rapid LC/MS methodology. Dried whole cell samples were suspended in water, sonicated, and an aliquot of the filtered solution was injected into the LC/MS. For each organism, there were a number of spectra for both low molecular weight components (<3000 Da) and higher molecular weight components that were unique to a given sample. A methodology was developed utilizing spectra for lower molecular weight protein biomarkers

(<30,000 Da) which makes it possible to perform the identification in real time (Figure 3). Electrospray mass spectra of proteins have a pattern of charge states which can be readily converted to a one dimensional Boolean array. A set of biomarkers for one or more organisms can be collected in a two dimensional Boolean array to create a biomarker library. During the analysis of the bacterial sample, a "logical AND" operation is performed which effectively matches the pattern of the spectrum to the library of biomarkers. If enough of the ions corresponding to the charge states of a biomarker are present, a match is identified and presented in real time to the user. At the end of the analysis, results for all of the collected spectra are displayed. The length of the bar in the final display is a measure of the area under the peak for a given biomarker.

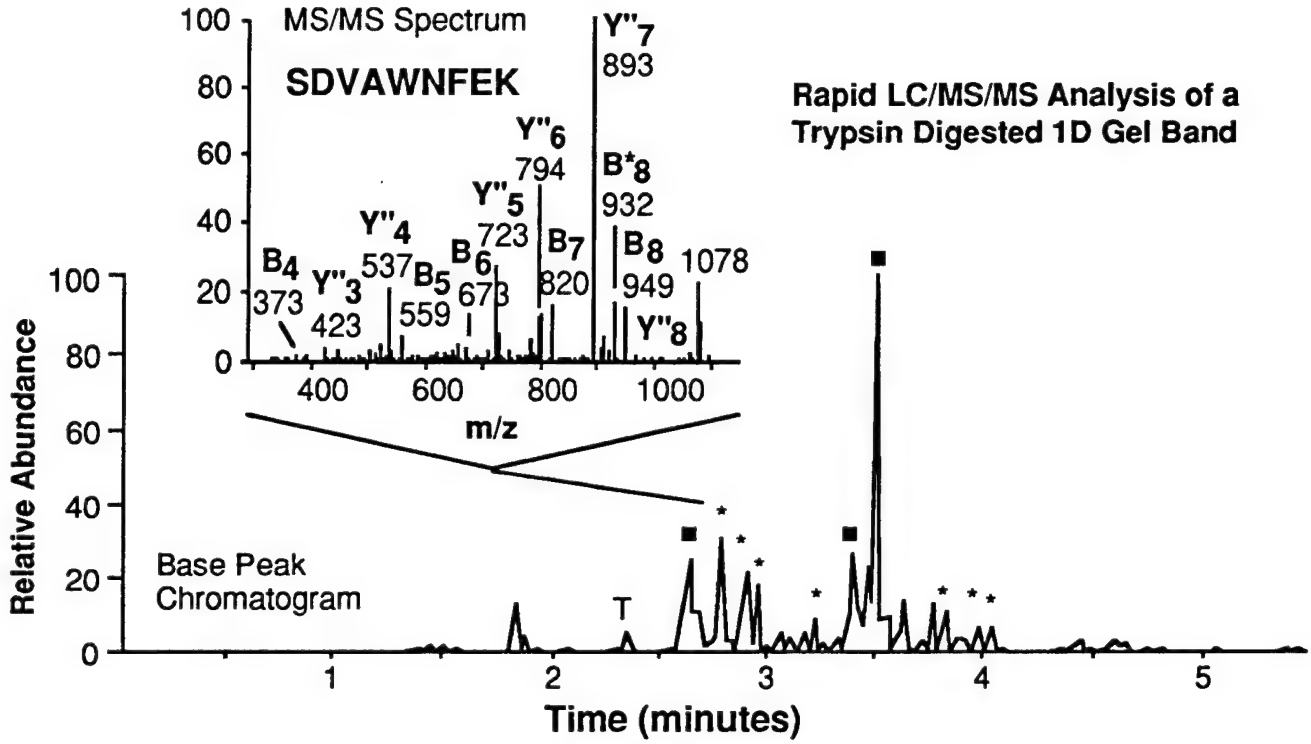


Figure 2. Rapid LC/MS/MS analysis of a tryptic digest mixture from human glutathione peroxidase. MS/MS spectra for the peaks labelled with an asterisk were correctly matched to peptides derived from human glutathione peroxidase using the SEQUEST database search program. Peptides labelled with a square were correlated manually once the identity of the protein was known. The peak labelled with a T is derived from trypsin. A representative MS/MS spectrum is shown in the inset.

This approach has proven to be very robust. Even noisy biomarker spectra can be correctly analyzed. If two biomarkers co-elute, both can be identified. Results from one set of experiments is given in Figure 4. When matched against a library of 32 biomarkers from 5 different organisms, each organism could be correctly identified when analyzed separately (data shown for only B. anthasis and B. meletensis, upper two panels Figure 4). When an mixture of B. anthrasis and B. meletensis was analyzed, both were identified (lower panel Figure 4).

CONCLUSIONS

Newly developed micro-scale electrospray ionization sources can be coupled to capillary liquid chromatography and ion trap mass spectrometers for rapid analysis of peptide and protein mixtures. Tandem mass spectrometry of the mixture of digested proteins yields a wealth of information which can be used to rapidly (within ten minutes) identify a protein. Using a novel method for analyzing mass spectra, electrospray spectra of bacterial proteins can be matched to a library of known protein

biomarkers. Each mass spectrum in an LC/MS run is analyzed as it is collected, and the data for the whole run is sufficient to identify bacterial samples when analyzed separately or as mixtures.

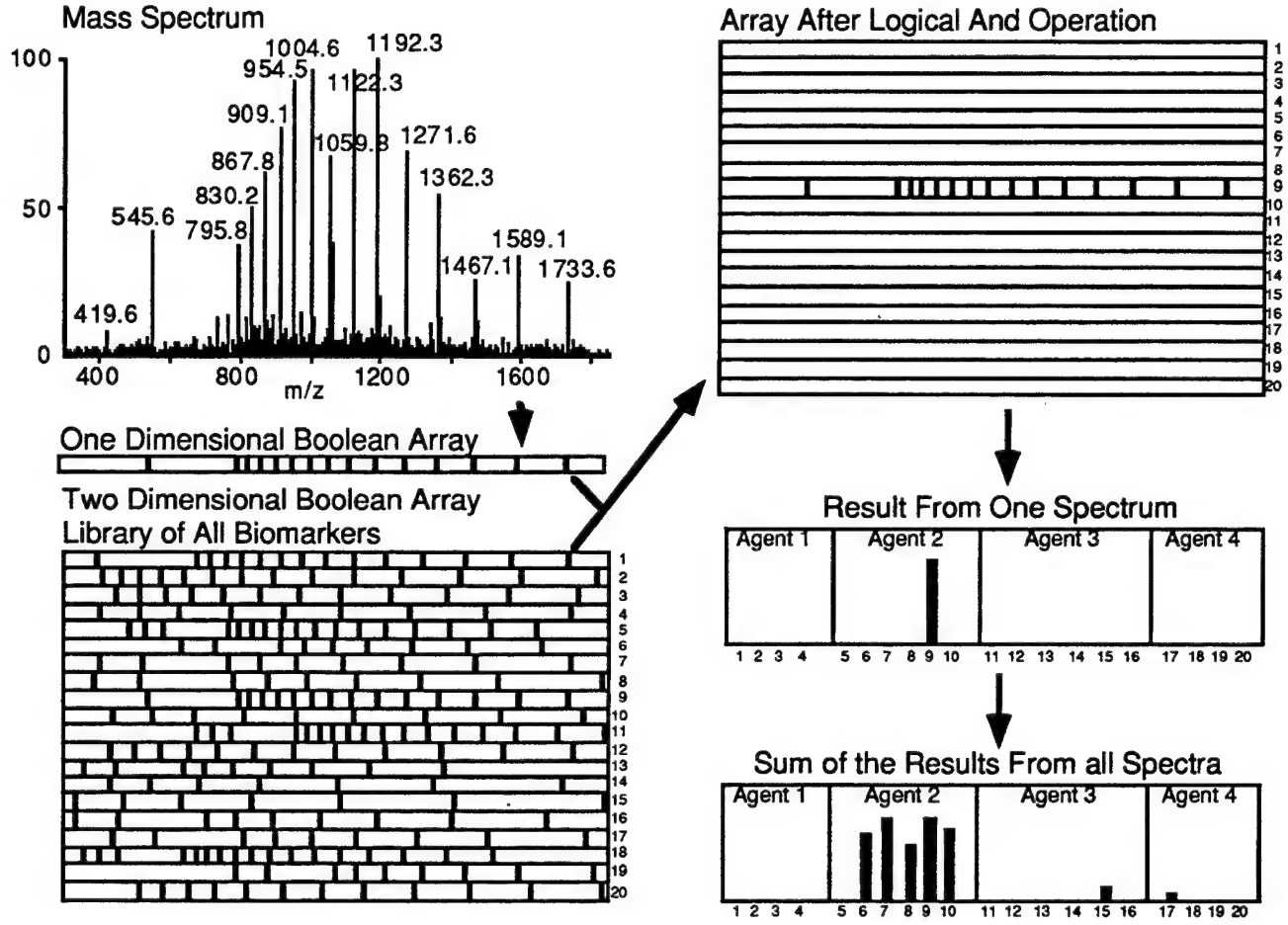


Figure 3. Schematic of the methodology used to identify protein biomarkers observed during the analysis of bacteria samples.

ACKNOWLEDGMENTS

This work is supported in part by grants from the National Institutes of Health (RR06217 and CA33572) and the U.S. Army. Dr. John R. Yates of the University of Washington kindly provided the SEQUEST database search program. The digested sample of human glutathione reductase was provided by Dr. Kristine Swiderek of the City of Hope.

REFERENCES

1. Wilm, M. S. and Mann, M., *Int J Mass Spectrom Ion Proc* **1994**, 136, 167-180.
2. Emmett, M. R. and Caprioli, R. M., *J Amer Soc Mass Spectrom* **1994**, 5, 605-613.
3. Davis, M. T., Stahl, D. C., Hefta, S. A. and Lee, T. D., *Anal Chem* **1995**, 67, 4549-4556.
4. Stahl, D. C., Swiderek, K. M., Davis, M. T. and Lee, T. D., *J Amer Soc Mass Spectrom* **1996**, 7, 532-540.
5. Davis, M. T., Stahl, D. C. and Lee, T. D., *J Amer Soc Mass Spectrom* **1995**, 6, 571-577.
6. Davis, M. T. and Lee, T. D., *J Am Soc Mass Spectrom* **1997**, in press.
7. Eng, J. K., Mccormack, A. L. and Yates, J. R., *J Amer Soc Mass Spectrom* **1994**, 5, 976-989.

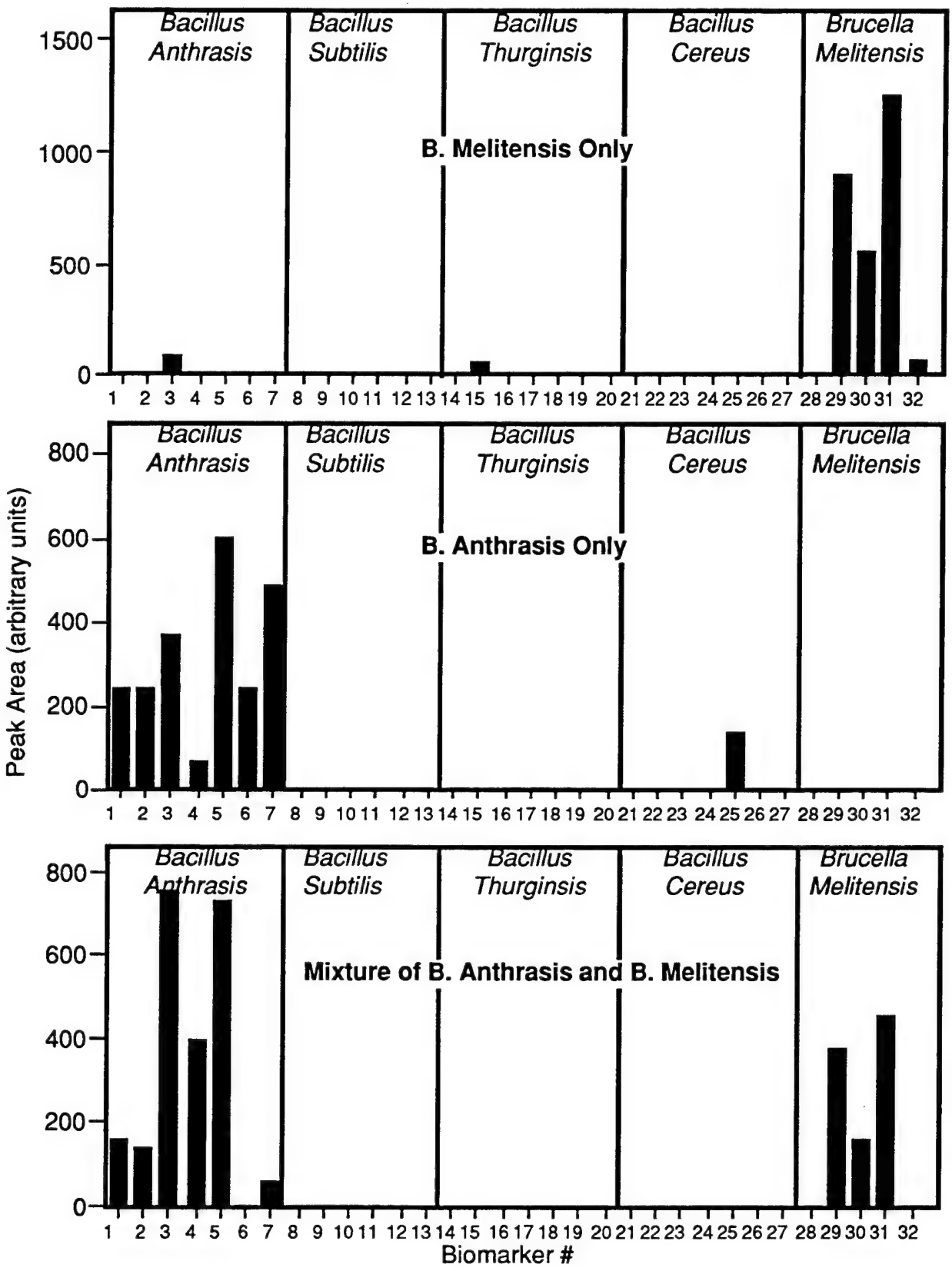


Figure 4. Results of the rapid LC/MS analysis of *Brucella Melitensis* and *Bacillus Anthrasis* separately (upper and middle panels), and as a mixture (bottom panel).

HIGH THROUGHPUT STRATEGY FOR IDENTIFICATION OF PROTEINS IN MIXTURES

John R. Yates, Andrew J. Link, David Schieltz, Lara Hays, Jimmy Eng
Department of Molecular Biotechnology
University of Washington
Seattle, Washington

ABSTRACT

Genomic and EST sequencing projects are providing a sequence infrastructure that will change how protein biochemistry is practiced. As the databases fill with sequences protein based discoveries will become more contextual- placing a protein's role or function within the context of a disease or pathway. Mass spectrometry has begun to play a key role in the connecting the genome and proteome. For example, integrated and automated methods for mapping proteins isolated by using 2-D gel electrophoresis have been developed. New methods based on the capabilities of tandem mass spectrometry can be used to identify novel protein targets, protein-protein interactions, protein complexes, and to rapidly survey the identities of proteins in subcellular organelles. Furthermore, a set of sequences or tandem mass spectra can also serve as biomarkers. Currently tandem mass spectra derived from microorganisms with a significant amount of sequence in the database can be readily identified through the SEQUEST database searching algorithm. A second strategy for microorganism identification requires the creation of a set of fingerprints based on tandem mass spectra and the software to match the tandem mass spectra. Software strategies are being developed to allow the construction of tandem mass spectral databases of peptides. This method will allow reference sets to be developed for biological experiments to aid in data analysis.

Blank

TRAPPING GENE PRODUCT: NEW TOOLS FOR STUDY PROTEINS

Jun Qin
Laboratory of Biophysical Chemistry
NHLBI/NIH
Bethesda, Maryland

ABSTRACT

A matrix-assisted laser desorption/ionization (MALDI) external ion source is coupled to an ion trap mass spectrometer (ITMS). An improved trapping method, an improved detection parameter set, and a newly devised excitation scheme for collision-induced dissociation of peptide ions have made the instrument possible to be used routinely to solve biological problems related to proteins.

The combination of matrix-assisted laser desorption/ionization and the ion trap provides an appropriate energy and time window for the observation of differential dissociation events in peptide ions derived from post-translationally modified proteins _ especially proteins that have undergone phosphorylation, glycosylation, or disulfide bond formation. Modified peptide ions dissociate readily, losing the modification moiety (or undergoing cleavage close to the modification), while keeping the peptide backbone intact in most cases. This kind of dissociation provides a convenient signature for the identification of the modification(s). Methods for disulfide mapping and for the identification of phosphorylated peptides are developed. These methods provide sensitive, accurate, and widely applicable means for elucidating post-translational modifications.

Strategies for identification of gel-separated proteins with MALDI/ITMS incorporating database searching will be discussed.

Blank

RECENT DEVELOPMENTS AND NOVEL APPLICATIONS IN MATRIX-ASSISTED LASER DESORPTION/IONIZATION MASS SPECTROMETRY (MALDI-MS)

Franz Hillenkamp
University of Muenster
Muenster, Germany

ABSTRACT

Over the last years MALDI-MS has become an established methode for the analysis of (biological) macromolecules. It is characterized by a very high sensitivity, a mass accuracy of 10-5 in the peptide and 10-4 in the protein mass range, has a limited selectivity in mixture analysis, a relative tolerance towards salts, buffers and detergents and a theoretically unlimited mass range. The presentation will give an introduction into the basic principles and technical approaches of matrix desorption and time-of-flight mass analysis. In the first part some new experiments and findings are presented which can help to better understand the basic MALDI mechanisms. Matrix selectivity and mass range limitation will be discussed in particular.

In the second part of the presentation instrumental alternatives to standard MALDI-TOF instruments will be discussed. This will include results for the desorption with lasers emitting in the infrared and a novel optical transmission geometry.

Two new fields of application will be introduced in the last part of the talk: MALDI of single biological cells and MALDI-MS of non-covalently bound complexes.

Blank

ANALYSIS OF SELECTED BIOLOGICAL WARFARE AGENT SIMULANTS BY AN ELECTROSPRAY IONIZATION MASS SPECTROMETER

Sunia Afzaal, Catherine C. Stacy, and Frank H. Laukien
Bruker *DALTONICS*™
Billerica, Massachusetts

ABSTRACT

The Department of Defense (DOD) has a high priority need for the fielding of new fast, integrated and automated technologies which can positively identify species and strains for BWA aerosols. Electrospray ionization (ESI) mass spectrometry has been used in the laboratory for the taxonomic characterization of various bacteria and their molecular components. Ion trap mass spectrometry in particular when coupled to ESI offers a powerful technique for the determination of species and strain specific biomarkers for BWA, which can be miniaturized in principle. The MS^n feature of this technology provides additional specificity in the analysis of complex biomixtures. Here, preliminary studies are presented in which a new miniaturized and highly integrated commercially available ESI ion-trap mass spectrometer with a m/z range of 6000 is used for the analysis of selected biological warfare agent (BWA) simulants. Minimal sample preparation was used for the ultimate adaptation of these techniques for the battlefield. The analytes included ovalbumin, a commonly used simulant for proteinacious bio-toxins and microbial samples such as heat killed Erwinia herbicola and Bacillus subtilis var. niger. Mass spectra of the these complex bio-samples are used to find distinguishing features for the different classes of biological simulants analyzed.

Blank

BACTERIAL PHOSPHOLIPIDS AND FATTY ACIDS ANALYZED BY ELECTROSPRAY TANDEM MASS SPECTROMETRY (ES MS-MS) PROVIDES CHEMOTAXONOMIC DIFFERENTIATION OF THE GENUS *BRUCELLA*[1]

Gavin E. Black, Michael W. Wensing, A. Peter Snyder
U.S. Army Edgewood Research, Development and Engineering Center
Aberdeen Proving Ground, Maryland

Karen S. Heroux
Geo-Centers, Inc.
Aberdeen Proving Ground, Maryland

ABSTRACT

Negative ion electrospray tandem mass spectrometry (ES MS-MS) was applied to the Genus *Brucella*. There are several different species of *Brucella*. ES MS-MS has been used to profile the phospholipids of *Brucellas abortus* and *B. melitensis*. Using ES MS-MS, phospholipids are analyzed in their intact form without derivatization. Extracts of *B. abortus* strains has almost the sem phospholipid profiles by ES MS, whereas extracts of *B. melitensis* strains had distinct phospholipid profiles by ES MS. The product ion mass spectra of deprotonated phospholipids provided the polar headgroup class and the masses of the fatty acid moieties. ES MS-MS showed differences between *B. abortus* and *B. melitensis*. Phosphatidylglycerols were the predominant phospholipid, but there were peaks which corresponded with that of phosphatidylethonalamines and phosphatidylcholines. ES MS-MS applied to glycerophospholipids for *Brucella* species has the potential to be a chemotaxonomic method for bacterial differentiation. An HPLC method is being developed to allow separation of all the different classes of phospholipids and enable them to be analyzed by ES MS-MS.

[1] Research Funded by: National Research Council Associateship Program.

Blank

IDENTIFICATION AND SEQUENCING OF AN MHC-ASSOCIATED PEPTIDE RELATED TO *TRYPANOSOMA CRUZI* INFECTION BY TANDEM MASS SPECTROMETRY

Adam H. Brockman and Ron C. Orlando

Complex Carbohydrate Research Center
University of Georgia
220 Riverbend Rd.
Athens, GA 30602

Rick L. Tarleton

Department of Cellular Biology
University of Georgia
Athens, GA 30602

Cell surface class I MHC proteins function to present peptides to thymus derived (T) lymphocytes. The 8-10 amino acid peptides can indicate intracellular infection and elicit cytolysis. Using RP-HPLC in conjunction with tandem mass spectrometry we have isolated and analyzed peptide extracted from the explanted hearts of *Trypanosoma cruzi* infected mice. The candidate peptide was originally selected by searching databases for the known K^b motif. Tandem mass spectrometric analysis of one peptide, VDYNFTIV, confirmed the sequence of the naturally presented epitope. These methods promise to become increasingly important to bioinformatics as genome sequencing projects continue.

Trypanosoma cruzi is a deadly human parasite which is a leading cause of heart disease and death among young adults in Latin America. The parasite is carried by the reduviid or "kissing" bug, which is a blood feeding arthropod endemic to Latin America. The parasites are carried in the gut of the bug, and then spread by blood contact with fecal material that is passed by transmitting vector species during feeding.

The expression of MHC-associated Peptides mediates the success or failure of the cytotoxic T lymphocyte (CTL) mediated aspect of immunity. Identification of those peptide epitopes which are prominently presented *in vivo* promises to be an exciting new avenue for vaccine development. Initial work in this area has been performed on *in vitro* cultured melanoma cells by such investigators as Cotter et al.[1], Tomlinson and Naylor et al.[2], and Hunt, Shabanowitz, and Englehard et al.[3]

Development of sensitive and selective biological assays using CTLs is the best means of identifying pools of extracted peptides which contain important epitopes. However, development of CTL assays of peptide fractions from infected tissue extracts can be complicated by the extreme diversity of the whole peptide mixture and the inability to stimulate with infected cells/tissue *in vitro*.

Mass spectrometry in conjunction with the increasing power of genomic database searching offers an efficient solution to problems associated with the search for relevant CTL epitopes. Suspect peptides can be detected using mass spectrometry to scan extracts for the

molecular weights of those peptides which are suggested by motif scanning data. Tandem Mass Spectrometry is then employed to obtain partial sequence information of the peptides. Suspect peptides can then be synthesized and tested by CTL assay without any of the interferences encountered in extremly dilute peptide pools from tissue extracts.

We have identified the peptide VDYNFTIV ("pep77.2") in infected tissue in this manner, having also known the peptide to be CTL active when stimulation was carried out using high concentrations of the synthetic peptide. Tandem mass spectrometry confirmed that the sequence of the naturally presented peptide was indeed the same as that predicted by CTL assay data generated using synthetic peptide (Figure 1).

CONCLUSIONS

It is possible to identify infection-related MHC-associated peptides without the direct benefit of CTL assay as a selective detection tool following fractionation. However, database searching, mass spectrometry, and CTL assay are of obvious complementarity in solving this difficult bioanalytical problem. The power of these methods together promise to contribute valuable information regarding various infectious disease systems as genome sequencing projects come to completion. This will be especially important to the field of bioinformatics - the process of determining new ways of locating biologically relevant gene products encoded within given genome. Our work is currently hindered by the lack of available sequence information for *T. cruzi*, and by the development of a more robust CTL assay for the biological system in question. However, as the *T. cruzi* genome sequencing projects come to completion and analytical methods continue to improve we expect to find and test many potential vaccine candidates for this devastating disease.

FIGURE 1
Tandem mass spectrum of the immunologically active peptide, pep77.2. Tandem mass spectra acquired from infected heart extract matched spectra generated from synthetic peptide.

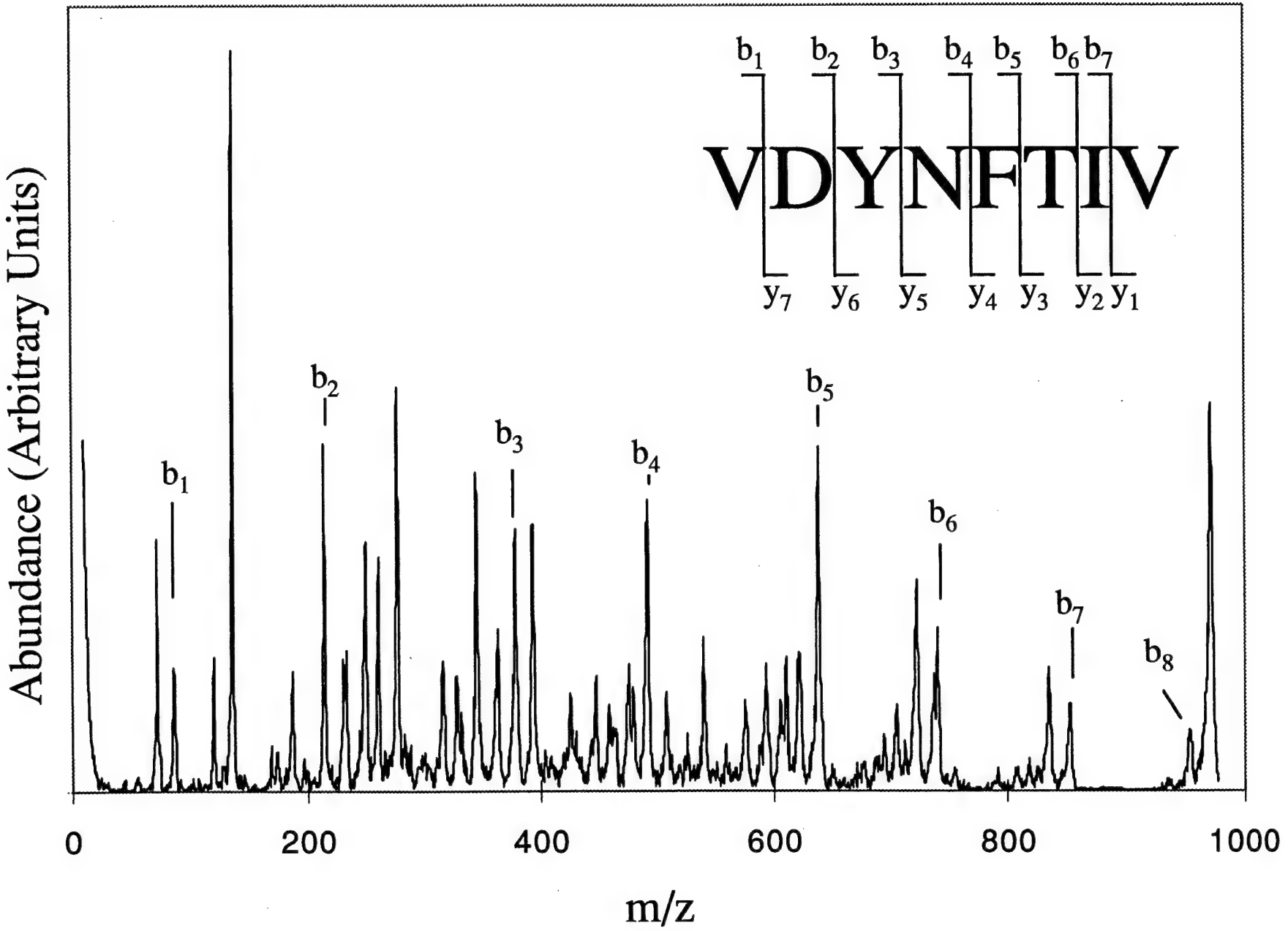
REFERENCES

1.Jafee, E. M.; Huang, A.; Pasternack, G.; Woods, A.; Cotter, R.; Pardoll, D.; J. *Cellular Biochemistry Supplement*, 17, **1993**.

2.Cox, A. L.; Skipper, J.; Chen, Y.; Henderson, R. A.; Darrow, T. L.; Shabanowitz, J.; Englehard, V. H.; Hunt, D. F.; Slinguff, C. L. *Science*, **1994**, 264, 716-719.

3.Tomlinson, A. J.; Jameson, S.; Naylor, S.; *Journal of Chromatography A*, **1996**, V744, N1-2, 273-278.

Figure 1: Tandem MS of P77.2



Blank

IDENTIFICATION OF POST-TRANSLATIONAL MODIFICATIONS BY THE COMBINED USE OF MASS SPECTROMETRY AND EDMAN SEQUENCING

Lydia Nuwaysir, J. Mark, Gary W. Kuroki, Matt McLean,
Steve Levery and Dave Yamane (presented by Charles Coonley)
PE Applied Biosystems
Foster City, California

ABSTRACT

The PE Applied Biosystems Model 173A MicroBlotter and ProciseTM Protein Sequencing System along with the PE Sciex API 100 Mass Spectrometer provide a complete system for the separation, collection and structural analysis of protein digests. The MicroBlotter is a capillary LC system with automated fraction collection directly onto a PVDF membrane. When modified with a post-column splitter a portion of the solvent flow was diverted to the API 100 for mass analysis. The mass spectrometry data was used to screen the digest for post-translationally modified peptides. Samples collected by the MicroBlotter were analyzed by chemical sequencing to determine amino acid sequence data and identify the modified site. This integrated approach maximizes structural information obtained from a single sample injection.

Blank

MONITORING PROTEIN EXPRESSION BY MALDI OF UNPURIFIED BACTERIAL CELLS

Michael L. Easterling, David Silverstein, Christopher M. Colangelo,
Robert A. Scott, and I. Jonathan Amster
University of Georgia
Athens, Georgia

ABSTRACT

Gel chromatography is the standard method of monitoring cellular contents during the protein expression in bacterial hosts such as E. coli. In this procedure, samples are periodically collected from the enriched medium and the cells are concentrated and lysed before analysis. We have developed a method of protein expression monitoring using matrix assisted laser desorption ionization (MALDI) of whole bacteria. This technique provides rapid sample preparation, allowing analysis of extracted bacteria in less than 10 minutes. MALDI analysis tolerates high salt concentrations, eliminating the need for preparative steps such as dialysis needed for other ionization techniques. Comparisons between mass spectrometry and gel chromatography will be presented in terms of increased efficiency, resolution, and mass accuracy. By using delayed extraction time-of-flight mass spectrometry for high resolution, more detailed analysis is presented for monitoring post-translational modifications that occur during expression of complex proteins.

Blank

COVALENT BINDING OF 2,4,6-TRINITROTOLUENE TO HUMAN HEMOGLOBIN. EVIDENCE FOR PROTEIN AND HEME ADDUCTS PROBED BY ELECTROSPRAY IONIZATION MASS SPECTROMETRY

R. Bakhtiar, K. H. Leung, M. A. Bednarek, R. A. Stearns, and C. E. C. A. Hop
Mail Stop: RY80L-109; Department of Drug Metabolism
Merck Research Laboratories; Rahway, New Jersey 07065-0900 USA

ABSTRACT

The bioactivation of nitroaromatic compounds to highly reactive intermediates is responsible for the genotoxic and cytotoxic effects by reaction with DNA and proteins. Number of reports have discussed their adverse biological effects including aplastic anemia, hemolysis, impairment of nervous system, cataracts, and liver toxicity. Evidence from studies of exposure to human and model animal studies have demonstrated the formation of hemoglobin (Hb) adducts upon inhalation, ingestion, or skin contact. Thus, the covalent binding of nitroaromatic compounds or their metabolites to critical cellular proteins (e.g., Hb) has been of interest. In the present work, i*n vitro* reaction of nitroaromatics with human hemoglobin under aerobic and anaerobic/reductive conditions is studied using liquid-chromatography/electrospray ionization mass spectrometry (LC/ESI-MS).

INTRODUCTION

Nitroaromatic explosives such as 2,4,6-trinitrotoluene (TNT) are found in many munitions formulations [1]. Explosive residues in soil and groundwater in close proximity to military sites have become an environmental concern [1]. The bioactivation of nitroaromatic compounds to highly reactive intermediates is responsible for their genotoxic and cytotoxic effects by reaction with DNA and/or proteins [2]. Due to its continued use as a secondary explosive and its prevalence at contaminated sites, the mechanism of covalent binding of TNT or its metabolites to critical cellular proteins has been of interest. A substantial body of information is available on the metabolic pathways of TNT and other nitroarenes. Several reports [3-5] have discussed the metabolic pathways of TNT and its adverse biological effects including aplastic anemia, hemolysis, impairment of nervous system, cataracts, and liver toxicity. The toxic effects of TNT in rats have been suggestive of morphological changes including swelling of cytoplasm and dilatation of Golgi apparatus [5]. The 2-electron reduction of nitro groups yielding nitrosoarene intermediates which react with the thiol (-SH) groups of cellular peptides (e.g., glutathione) and proteins have been documented (Figure 1) [3]. Furthermore, the formation of a semimercaptal and subsequent rearrangement to the corresponding sulfinamide has been proposed (Figure 1) [3].

Macromolecule adducts have been proposed as molecular dosimeters as an indicator of human exposure to certain chemicals [6,7]. Evidence obtained in studies on human and model laboratory animals have demonstrated the formation of hemoglobin (Hb) adducts upon inhalation, ingestion, or skin contact with TNT [4]. In this regard, Hb adducts have been found to be invaluable biomarkers for toxicity and carcinogenesis in molecular epidemiological studies. Fortunately, recent advancements in analytical chemistry have made it possible to quantify Hb adducts using a combination of separation and mass spectrometry techniques [8]. The goal of these studies was to monitor human exposure to chemical carcinogens from environmental and occupational sources and to predict health risk assessment. Despite evidence concerning the covalent modification of thiol containing proteins and peptides by TNT (*vide supra*), there has been no report on modification of its heme moiety. Herein, we present electrospray ionization mass spectrometry (ESI-MS) data which suggest the generation of a novel heme adduct upon *in vitro* reaction of human Hb with TNT under *anaerobic and reductive* conditions. In addition, detection of covalently modified human Hb polypeptide chains and identification of the reactive amino acid residue in the β-chain using tryptic digest and tandem mass spectrometry (MS/MS) will be briefly discussed.

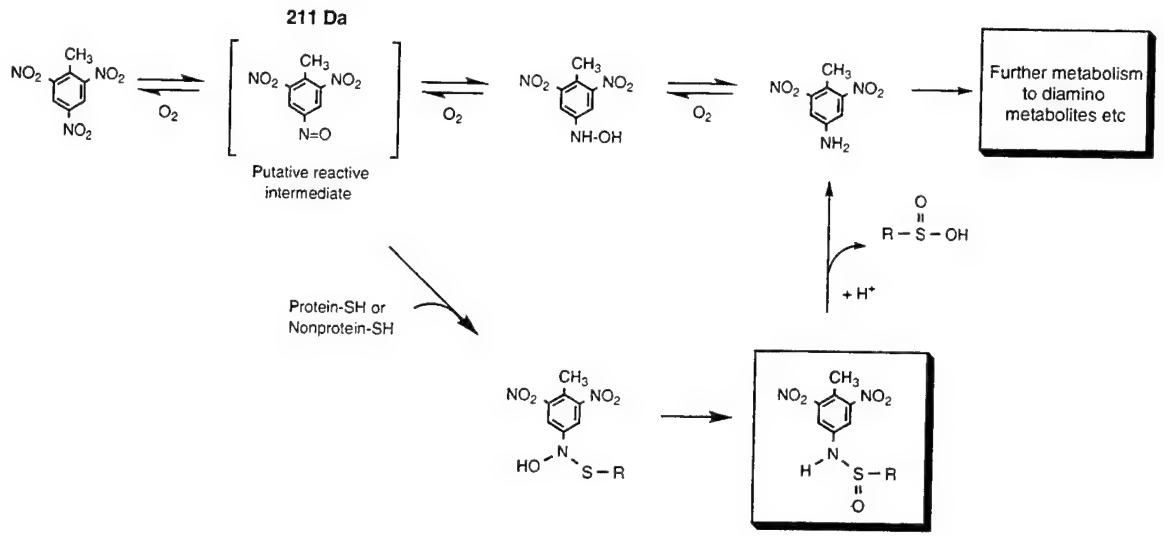


Figure 1

EXPERIMENTAL

Human Hb was purchased from Sigma Chemical Co. (St. Louis, MO). Sodium hydrosulfite was purchased from Aldrich Chemical Co. (Milwaukee, WI). Ammonium acetate and all high purity (HPLC grade) solvents were obtained from Fisher Scientific (Fairlawn, NJ). 2,4,6-Trinitrotoluene (TNT) was obtained from Eastman Kodak (Rochester, NY) and was recrystallized from 95% ethanol. The purity of TNT was examined using high-pressure liquid chromatography (HPLC) and an ultra-violet visible (UV-VIS) detector at 230 nm.

Three stock solutions were prepared: (i) TNT (3 mM) in anhydrous methanol; (ii) Hb (4 mg/mL) in a 10 mM ammonium acetate buffer (pH=7.1); and (iii) sodium hydrosulfite (1.36 mg/mL) in de-ionized water. Hb incubations were made under three different conditions. Three different Warburg tubes were assembled on a vacuum line for sample preparation. Tubes 1 and 2 were purged under vacuum using high purity argon gas for several minutes. Tube 3 was left under atmospheric condition. A 0.25 mL aliquot of the Hb (about 62.5 nmoles of heme equivalent) solution was added to each Warburg tube. Tubes 1 and 2 underwent seven freeze-pump-thaw cycles. 10 Microliters of the sodium hydrosulfite solution were injected through a rubber septum to tube 1 and the tube was shaken for several minutes. The total volume of the incubates was brought up to 1.2 mL using 10 mm ammonium acetate buffer (pH 7.1). Next, 100 microliters of the TNT solution was added to all three tubes. For the sake of clarity, conditions for the aforementioned incubations are summarized in Table 1. All tubes were shaken for several minutes, slowly brought to atmospheric pressure with argon gas (tubes 1 and 2), and incubated at 37°C for 6 hours in dark. In order to minimize ESI signal suppression by sodium hydrosulfite, about 1.5 mL of anhydrous methanol was added to each incubate. The supernatant and precipitate were separated by centrifugation at 3000 g for 10 minutes. The filtrates were dried in a SAVANT SpeedVac SVC100 (Farmingdale , NY). The dried samples were reconstituted in anhydrous methanol or dimethylformamide (in order to enhance solubility) and analyzed by mass spectrometry.

Table 1. Summary of reactions conditions described in the text.

	TUBE 1	TUBE 2	TUBE 3
Hb	0.25 mL	0.25 mL	0.25 mL
Sodium Hydrosulfite	10 mL	----------	----------
TNT	100 mL	100 mL	100 mL
Environment	Argon	Argon	Air

All samples were prepared in duplicates.

An API III$^+$ triple quadrupole mass spectrometer (PE-Sciex, Thornhill, Onatrio, Canada) with an ion spray interface was employed [9]. Samples were delivered *via* a 50 mm i.d. fused-silica capillary to the ion spray tip which was held at potential of +4.8 kV.

Peptides were analyzed using a Model 140B dual syringe solvent delivery system (Applied Biosystems, Foster City, CA). A 2.1x150 mm, Vydac C-18 (5 micron particles) column was utilized. LC pumps were operated at a flow rate of 0.25 mL/min. Post-column splitting was approximately 1:5 resulting in a flow rate of 50 μL/min directed to the mass spectrometer. Elution of peptides obtained from tryptic digests of Hb was accomplished by a gradient as follows: 0-5 min 100%A; 0% B to 100% B in 60 min; 100% B for 5 min. Solvent A consisted of 1900 mL water, 100 mL acetonitrile. Solvent B consisted of 1900 mL acetonitrile, 100 mL water. Injections were made using a Rheodyne six-port injector equipped with a 10 μL loop.

RESULTS & DISCUSSION

Hb is a tetrameric protein composed of two α- and two β-subunits, each consisting of 141 and 146 amino acid residues, respectively. Human Hb contains four heme prosthetic groups which are embedded within the tertiary structure of the protein *via* specific non-covalent interactions. The heme found in Hb (iron protoporphyrin-IX or heme *b*) contains two propionic acid, two vinyl, and four methyl side-chain groups attached to the pyrrole rings of the porphyrin structure. Generally, heme is in the iron (II) high-spin state in deoxy-Hb and has a total of six coordinates. The four in-plane positions are occupied by the nitrogens in the four pyrrole groups of the porphyrin ring. The 5th axial coordination site is occupied by an imidazole ligand from a proximal histidine which is also referred to as the His F8 [10]. The 6th axial position is normally occupied by an oxygen molecule. Subsequent to ligation of the 6th position by an oxygen molecule, the ligand field is strong enough to yield spin-pairing, resulting in a low-spin d^6 (t^6_{2g}) complex. During the course of our investigation, we observed a putative heme adduct using ESI-MS and further scrutinized it by tandem mass spectrometry (MS/MS). Figure 2a depicts the

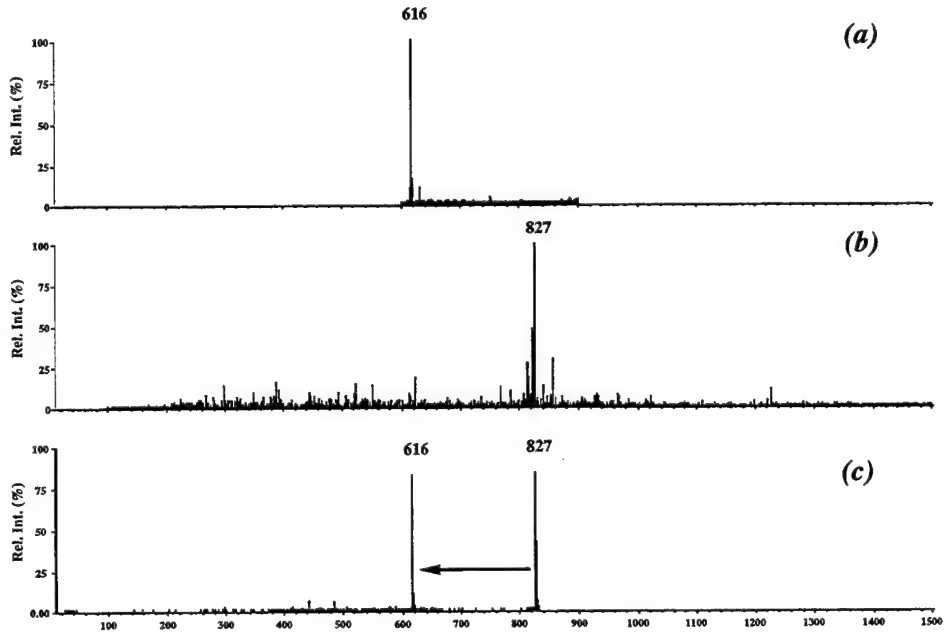


Figure 2

ESI-MS of an authentic sample of heme where the signal at *m/z* 616 corresponds to its radical cation. The mass spectrum of the putative heme-adduct generated in tube 1 is shown in Figure 2b yielding a signal at *m/z* 827. The mass increase of 211 Da can be attributed to addition of a nitrosodinitrotoluene molecule. The collision-induced dissociation (CID) spectrum of this adduct is shown in Figure 2c. Exclusive loss of 211 Da yielding a signal at *m/z* 616 was observed. CID of an authentic sample of heme (data not shown) results in sequential loss of 59 Da ($-CH_2COOH$), presumably from the two propionic acid side-chain, giving signals at *m/z* 557 and 498. Moreover, we learned that this adduct is acid labile and undergoes exclusive release of the heme moiety under acidic conditions. For example, the ESI-MS spectrum of a 5% acetic acid treated solution yielded exclusively a signal at *m/z* 616. To our knowledge, this is the *first* example of a heme-TNT adduct generated by the reaction of TNT with human Hb *in vitro*. Incubations made in tubes 2 and 3 (Table 1) did not appear to generate any discernible amount of adduct. Presumably, reduction of the nitro group by sodium hydrosulfite (Figure 1) is an obligatory step in the formation of the adduct. In addition, incubation in a Warburg tube containing sodium hydrosulfite (similar to tube 1) but under atmospheric conditions did not lead to any detectable adduct formation under our experimental setting. While the direct 2-electron reduction of the nitro group occurs under reductive conditions; the nitroso intermediate could undergo oxidation at high molecular oxygen concentrations to yield the parent nitro compound [11,12].

In TNT's metabolic activation, the initial step in the reduction of a nitro to a nitroso intermediate is presumably performed by enzymes such as NADPH cytochrome P450 reductase [11,12]. For example, orally administered nitroarenes can undergo nitroreduction in the intestinal microflora by bacterium that exhibit nitroreductase activity [12]. Indeed , NADPH cytochrome P450 reductase readily reduced TNT to 4-hydroxylamino-2,6-dinitrotoluene [11]. Herein, we attempted to mimic the reductive conditions using sodium hydrosulfite. Furthermore, it was observed that higher ratios of sodium hydrosulfite to heme (20% in excess) did not improve the adduct's yield during incubation.

Recently, Richter-Addo and co-workers [13] reported the structural characterization of a synthetic η^1-N-nitrosobenzene-iron-tetraphenylporphyrinato (tpp) dianion adduct ($[Fe^{II}(tpp)(C_6H_5NO)_2]$). Their X-ray crystallography data suggested an overlap of the HOMO of the d^6 low-spin ($S = 0$) Fe(II), namely the degenerate d_{xz} and d_{yz} orbitals, and the π^* orbital of the C_6H_5NO ligand. In qualitative terms, the d_{xy} is nonbonding while the d_z^2 and $d_{x^2-y^2}$ are the antibonding orbitals. Other examples involving iron-chelation by NO containing species have been observed, including 2-electron oxidation of hydroxylamine to its corresponding nitroso moiety or similar products obtained from a nitro functionality under reductive conditions [14].

In spite of the limited structural information obtained by the MS/MS spectrum (Figure 2c), it is appealing to *tentatively* invoke a Fe-N linkage *via* a putative 4-nitroso, 2,6-dinitrotoluene moiety. This speculation is further based upon the finding that the 4-(N-hydroxyamine)-2,6-dinitrotoluene was the major form of the TNT metabolite in the Hb samples obtained from workers exposed to TNT *via* skin exposure or inhalation in a munition factory [4,15]. Nonetheless, reduction of any of the nitro groups could lead to formation of a reactive intermediate in a similar fashion (Figure 1) [3,11]. Notably, in the study reported by Richter-Addo and co-workers [13], a 1:2 ratio ($[Fe^{II}(tpp)(C_6H_5NO)_2]$) between the $Fe^{II}(tpp)$ and C_6H_5NO was observed, respectively. However, according to spectra shown in Figure 2, a 1:1 ratio between the heme and the putative nitrosodinitrotoluene was evident. This difference could be partly rationalized in terms of the local environment of the heme within the Hb tertiary structure. In Hb, the heme's 5th coordination position is occupied by the His F8 (87th residue in the α-chain) within the helical chains. Therefore, only the 6th coordination site will be feasible for any potential coordination.

The covalent modification of Hb polypeptide chains by TNT has been reported by Sabbioni *et al.* using gas-chromatography mass spectrometry as well as an enzyme-linked immunoassay (ELISA) [15]. We have further confirmed their observation using ESI-MS and tandem mass spectrometry of peptides

obtained from the enzymatic digestion by trypsin. Human Hb consists of one Cys in the α-chain at position 104 and two Cys residues in the β-chain at positions 93 and 112. The deconvoluted ESI-MS spectrum of the α- and β-subunits revealed measured values of 15126.4±1.6 Da (calculated from sequence 15127.4 Da) and 15867.2±1.8 Da (calculated from sequence 15868.3 Da), respectively. The deconvoluted ESI-MS spectrum of the human Hb incubated with TNT (Figure 3), under *anaerobic and*

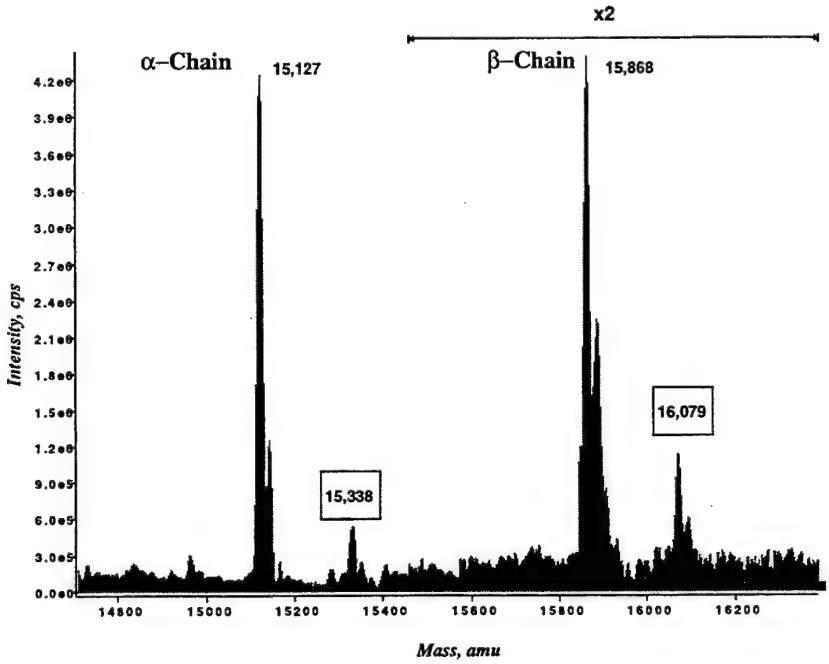


Figure 3

reductive conditions, revealed a modification of both subunits corresponding to a mass increase of ca. 211 Da at a 1:1 stoichiometry for each globin chain. An estimated amount of globin adduction was based on the mass spectral relative signal intensities subsequent to deconvolution yielding a relative abundance (RA) of ca. 7% and 18% for signals at 15337.8±1.6 Da and 16078.7±1.8 Da, respectively. Signals corresponding to the sodiated form (presumably from the residual sodium hydrosulfite) of the globin chains were also evident. We propose that the monoadduction with the molecular weight (MW) of 15337.8 is the result of modification of the only Cys residue at position 104 in the α-globin chain by a dinitronitrosotoluene (MW=211 Da) intermediate. The lower relative abundance of this signal could be in part attributed to the lack of accessibility of the Cys 104 which is buried in the protein tertiary structure. The location of the monoadduction on the β-chain was identified using trypsin digestion of a chromatographically isolated β-chain sample and subsequent LC-MS/MS of the resulting modified peptide, ^{83}GTFATLSELHCDK^{95}. Therefore, our preliminary data revealed that the predominant modification is located at the reactive cysteine-93 of the β-chain. In the absence of steric constraints, the sulfhydryl of the cysteine in position 93 has been shown to exhibit higher reactivity [16]. Both human and guinea pig Hb have this reactive cysteine [16].

CONCLUSIONS

Although the covalent modification of Hb polypeptide chains by TNT has been established [4,15], this study represents a unique type of adduct involving the heme prosthetic group. This putative heme adduct is acid labile and undergoes exclusive release of the heme moiety at or below pH ~ 5. In spite of technical difficulties inherent in quantitation of heme adducts [14], the sensitivity of ESI-MS and MS/MS experiments aided in obtaining qualitative structural information. During the course of our

study, we learned that the incubation and isolation of the putative heme adduct (*vide supra*) can be extremely sensitive to stoichiometry of the reactants, pH, temperature, light, and the choice of solvent. Initial attempts to synthesize and isolate sufficient amount of this adduct for NMR studies were unsuccessful due to the low product yield. Nonetheless, further structural analysis by NMR and/or X-ray crystallography is warranted. It is known that heme adducts play a critical role in irreversible enzyme inactivation and protein denaturation [17]. This finding could be of relevance in the investigation of biotransformation of TNT in subjects exposed to TNT *via* skin exposure or inhalation. The toxicological ramifications of the heme-TNT complex requires further investigations.

ACKNOWLEDGMENTS

We thank Drs. Lee Chiu and A. Y. Lu for helpful discussions.

REFERENCES

1. (a) Shriver-Lake, L.C.; Donner, B. L.; Ligler, F. S. *Environ. Sci. Technol.* **1997**, *31*, 837-841.

2. (a) Sabbioni, G. *Environ. Health Perspect.* **1994**, *102*, 61-67.; (b) Sabbioni, G. *Chem. Res. Toxicol.* **1994**, *7*, 267-274.

3. (a) Savolainen, H.; Tenhunen, R.; Harkonen, H. *British J. of Industrial Medicine* **1985**, *23*, 354-355.; (b) Liu, Y-Y.; Lu, A. Y. H.; Stearns, R. A.; Chiu, S-H. L. *Chem. Biol. Interactions* **1992**, *82*, 1-19.; (c) Leung, K. H.; Yao, M.; Stearns, R. A.; Chiu, S-H. L. *Chem. Biol. Interactions* **1995**, *97*, 37-51.

4. Liu, Y-Y.; Yao, M.; Fang, J-L.; Wang, Y-W. *Toxicol. Letters* **1995**, *77*, 281-287 and references cited therein.

5. Zitting, A.; Szumanska, G.; Nickels, J.; Savolainen, H. *Arch. Toxicol.* **1982**, *51*, 53-64.

6. (a) Chang, L. W.; Hsia, S. M. T.; Chan, P-C.; Hsieh, L-L. *Annu. Rev. Pharmacol. Toxicol.* **1994**, *34*, 41-67.; (b) Bergmark, E. *Chem. Res. Toxicol.* **1997**, *10*, 78-84.; (c) Rappaport, S. M.; Symanski, E.; Yager, J. W.; Kupper, L. L. *Environ. Health Perspect.* **1995**, *103*, 49-53.

7. (a) Skipper, P. L.; Peng, X.; Soohoo, C. K.; Tannenbaum, S. R. *Drug Metabolism Rev.* **1994**, *26*, 111-124.; (b) Skipper, P. L. *Chem. Res. Toxicol.* **1996**, *9*, 918-923.

8. (a) Ferranti, P.; Sannolo, N.; Mamone, G.; Fiume, I.; Carbone, V.; Tornqvist, M.; Bergman, A.; Malorni, A. *Carcinogenesis* **1996**, *17*, 2661-2671.; (b) Kaur, S.; Hollander, D.; Haas, R.; Burlingame, A. L. *J. Biol. Chem.* **1989**, *264*, 16981-16984.; (d) Noort, D.; Hulst, A. G.; Trap, H. C.; de Jong, L. P. A.; Benschop, H. P. *Arch Toxicol.* **1997**, *71*, 171-178.; (e) Springer, D. L.; Bull, R. J.; Goheen, S. C.; Sylvester, D. M.; Edmonds, C. G. *J. Toxicol. and Environ. Health* **1993**, *40*, 161-176.; (f) Black, R. M.; Harrison, J. M.; Read, R. W. *Xenobiotica* **1997**, *27*, 11-32.; (g) Day, B. W.; Jin, R.; Karol, M. H. *Chem. Toxicol. Res.* **1996**, *9*, 568-573.

9. Li, Y-T.; Hsieh, Y-L.; Henion, J. D.; Ganem, B. *J. Am. Soc. Mass Spectrom.* **1993**, *4*, 631-637.

10. Ochiai, E. I. *J. Chem. Educ.* **1997**, *74*, 348-354.

11. Leung, K. H.; Yao, M.; Stearns, R. A.; Chiu, S-H. L. *Chem. Biol. Interactions* **1995**, *97*, 37-51.

12. Rickert, D. E. *Drug Metabolism Rev.* **1987**, *18*, 23-53.

13. Wang, L-S.; Chen, L.; Khan, M. A.; Richter-Addo, G. B. *Chem. Soc. Chem. Commun.* **1996**, 323-324.

14. Ortiz de Montellano, P. R.; Correia M. A. In *Cytochrome P450: Structure, Mechanism, and Biochemistry;* Ortiz de Montellano, P. R., Ed.; 2nd Edition, Plenum Press, New York, 1995, Chapter 9.

15. Sabbioni, G.; Wei, J.; Liu, Y-Y. *J. Chromatogr.* **1996**, *682*, 243-248.

16. (a) Olson, J. S. *Methods in Enzymology* **1981**, *76*, 631-651.; (b) Erve, J. C. L.; Deinzer, M. L.; Reed, D. J. *J. Toxicol. Environ. Health* **1996**, *49*, 127-143.

17. (a) Ortiz de Montellano, P. R. *Biochimie* **1995**, *77*, 581-593 and references cited therein.; (b) Swanson, B. A.; Ortiz de Montellano, P. R. *J. Am. Chem. Soc.* **1991**, *113*, 8146-8153.; (c) Kunze, K. L.; Ortiz de Montellano, P. R. *J. Am. Chem. Soc.* **1983**, *105*, 1380-1381.; (d) Ortiz de Montellano, P. R.; Grab, L. A. *J. Am. Chem. Soc.* **1986**, *108*, 5584-5589.; (e) Grab, L. A.; Swanson, B. A.; Ortiz de Montellano, P. R. *Biochemistry* **1988**, *27*, 4805-4814.

COVALENT BINDING OF ACYL GLUCURONIDES TO PEPTIDES AND PROTEINS

Cornelis E.C.A. Hop, Kwan H. Leung, Ray Bakhtiar and Matthew P. Braun
Merck Research Laboratories,
Mail Code: RY80L-109, P.O. Box 2000, Rahway, New Jersey 07065-0900

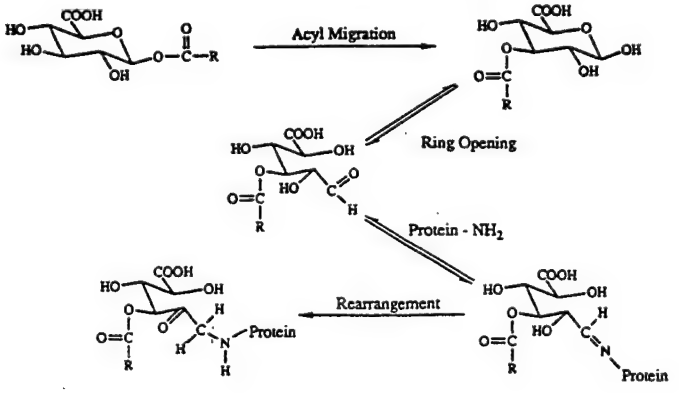
ABSTRACT

Ionspray LC-MS data are presented which show that the acyl glucuronides of the non-steroidal anti inflammatory drugs zomepirac, tolmetin and diclofenac bind covalently with the model peptide β-endorphin. Adducts are formed via two mechanism: one involving an imine intermediate and the other involving nucleophilic displacement of the glucuronic acid. Ionspray MS was also used to detect the intact adducts of the acyl glucuronides with serum albumin.

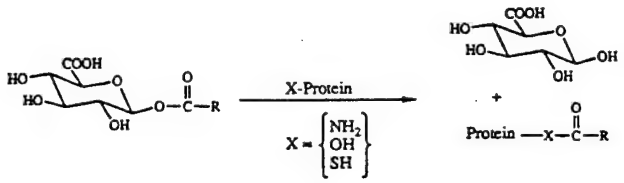
1. INTRODUCTION

Of crucial importance in biochemical toxicology is the interaction between biopolymers and xenobiotics, such as drugs and environmental contaminants. Non-covalent interaction is common and frequently the mechanism by which drugs exert their physiological effects. However, covalent modifications of biopolymers (peptides, proteins, DNA) can render these compounds ineffective and in some cases the adducts can be harmful. In this study, electrospray ionization mass spectrometry will be used to examine the interaction between metabolites of certain non-steroidal anti-inflammatory drugs (NSAIDs) and peptides/proteins. Drugs which contain an acid group can be metabolized by UDP-glucuronosyl transferase to an acyl glucuronide and the covalent adducts of the latter compounds with proteins have been implicated in mediating the cellular toxicity observed for certain NSAIDs in humans.[1] The exact nature of the mechanisms underlying NSAID-induced hepatotoxicity are unknown, but the hypothesis behind the toxicity is that covalent adducts are generated which cause a direct cytotoxic effect due to disruption of critical enzyme function or which act through an indirect mechanism involving the activation of the immune system giving rise to allergic (anaphylactic) reactions or allergy-type (anaphylactoid) hypersensitivity. Although there is evidence which suggests that the hepatotoxicity of the NSAID diclofenac is immune-mediated, a recent clinical evaluation suggested the mechanism to be due to "metabolic idiosyncrasy", because there were few signs of hypersensitivity in affected patient. The acyl glucuronides of the NSAIDs zomepirac, tolmetin and diclofenac are known to react with proteins, including serum albumin.[2,3,4] According to the literature [5] there are two mechanisms giving rise to covalent adducts:

a. the imine mechanism; the NSAIDs are linked via glucuronic acid to a lysine residue in the peptide/protein



b. the nucleophilic displacement mechanism; nucleophilic displacement of the glucuronic acid generates adducts containing NSAIDs linked to a lysine, serine or arginine residue in the peptide/protein



The extent of covalent adduct formation by the acyl glucuronides is controlled by (1) the *in vivo* concentration of the acyl glucuronides in certain tissues, (2) the intrinsic reactivity of the acyl glucuronides with peptides/proteins and (3) the *in vivo* stability of the covalent adducts. Finally, the function of the covalently modified peptides/proteins is a critical parameter in determining the severity of covalent adduct formation. Adducts of diclofenac metabolites with 110, 140 and 200 kDa proteins in mice liver have been observed.[4] Recently, the 110 kDa protein has been identified as the membrane enzyme dipeptidyl peptidase.[6]

Benet and Burlingame [7] examined covalent adduct formation by incubating serum albumin with one of the acyl glucuronides, followed by reducing, alkylating and digesting the modified serum albumin. Analysis of isolated tryptic fragments by off-line MS and MS/MS revealed the occurrence/nature/site of covalent binding. For example, the acyl glucuronide of tolmetin was found to react with lysine residues (mechanism a and b) and serine and arginine residues (mechanism b).[7b] However, this approach is rather time consuming and it is not inconceivable that enzymatic action can eliminate the covalently bound drug moiety. Therefore, we performed a series of experiments with a relatively small model peptide (β-endorphin), which can be studied by LC-MS and LC-MS/MS without the need for enzymatic digestion. The molecular weight of β-endorphin is 3465 Da and it contains five lysine residues. Ionspray mass spectrometry was also used to detect the intact adducts of the acyl glucuronides with serum albumin.

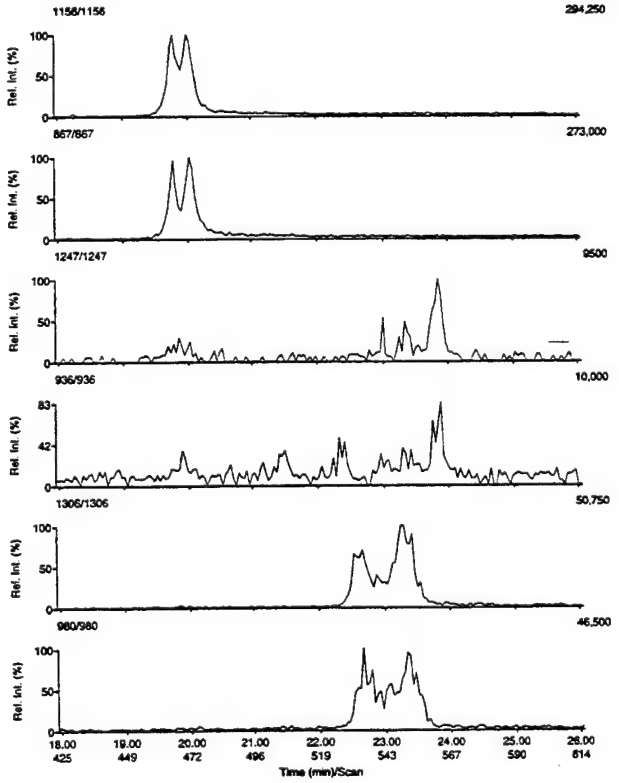
2. EXPERIMENTAL

The acyl glucuronides of zomepirac, tolmetin and diclofenac were synthesized in-house (see reference 8 for more details) and the peptides and proteins were purchased from Sigma. The incubations (400 μM substrate and 80 μM β-endorphin or bovine serum albumin) were performed at 37°C for 16 hours; the pH of the solution was 7.4 (phosphate buffer). Excess $NaBH_3CN$ was present for certain incubations (see below). A PE Sciex API III^{plus} triple quadrupole mass spectrometer equipped with an ionspray interface was employed for all experiments. For the β-endorphin experiments the covalent adducts were chromatographically separated from unmodified β-endorphin using a Zorbax Rx-C8 2.1 x 150 mm HPLC column; the flow rate was 200 $\mu l.min^{-1}$ and 40 $\mu l.min^{-1}$ was diverted to the ionspray source. Serum albumin and its adducts were not subjected to chromatographic separation.

3. RESULTS AND DISCUSSION

Using the reaction schemes presented above the m/z values of the adducts of the acyl glucuronide of zomepirac, tolmetin and diclofenac with β-endorphin can be calculated. LC-MS experiments showed that signals were visible at the anticipated m/z values. Figure 1 show data for the incubation of β-endorphin with zomepirac. The signals at m/z 1306 and 980 correspond to the 3+ and 4+ charge states of the adducts generated by the imine mechanism and the signals at m/z 1247 and 936 correspond to the 3+ and 4+ charge states of the adducts generated by nucleophilic displacement. It is clear that adducts were generated by the imine as well as the nucleophilic displacement mechanism with the adducts generated by the former mechanism being more abundant. In addition, multiple adducts are observed for each mechanism, which is not surprising considering that β-endorphin has 5 lysine residues. The mass spectra of the imine and nucleophilic displacement mechanism adducts generated during incubation of the acyl glucuronide of zomepirac with β-endorphin are shown in Figure 2 and 3, respectively. It is known that non-covalent adducts can be observed as well under electrospray conditions. However, there are several reasons which indicate that the adducts observed here are covalent in nature: (a) no adducts were observed for the t = 0 hrs incubation, (b) the m/z values of the observed signals correspond exactly with those calculated for covalent adduct formation (the mass of non-covalent adducts is 18 Da higher, because

Figure 1. The extracted ion current chromatograms for the β-endorphin + zomepirac acyl glucuronide incubation. The signals at m/z 1156 and 867 correspond with β-endorphin; m/z 1247 and 936 correspond with the adduct generated by nucleophilic displacement; m/z 1306 and 980 correspond with the adducts generated by the imine mechanism.



covalent adduct formation is accompanied by H_2O loss) and (c) the adducts do not co-elute with β-endorphin and the acyl glucuronides of zomepirac, tolmetin and diclofenac. Without chromatography signals can be observed in the mass spectra which correspond with non-covalent adducts.

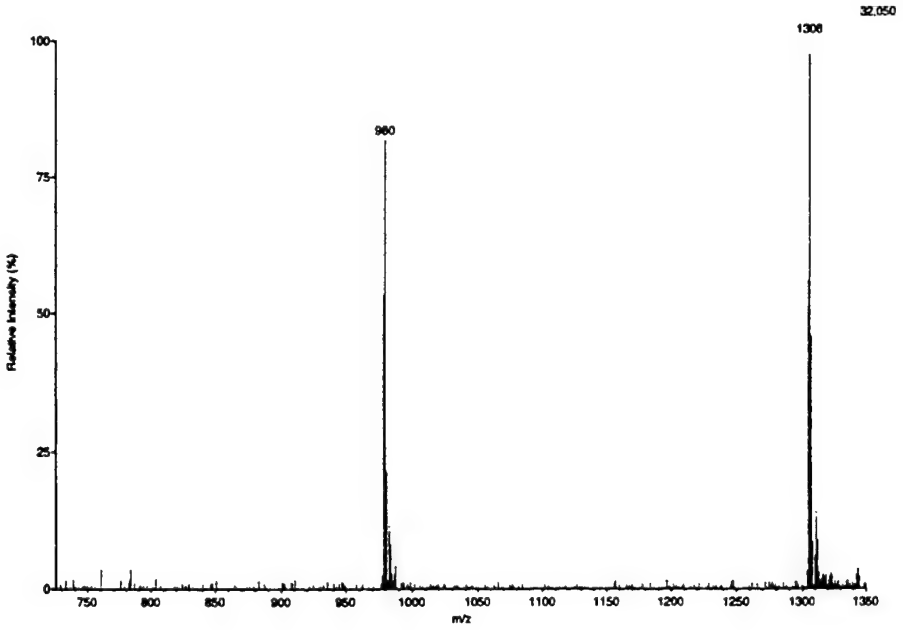


Figure 2. Mass spectrum of the imine mechanism adduct eluting at t = 23.3 min.

Incubations were also performed in the presence of $NaBH_3CN$ to reduce and, therefore, trap the unstable imine intermediates. Although the adduct formation by the nucleophilic displacement mechanism is relatively unaffected, the adduct formation by the nucleophilic displacement mechanism increases about five-fold. In addition, accurate mass measurements obtained by scanning slowly over a relatively narrow mass window showed that the molecular weight of the imine mechanism adducts generated in the presence of $NaBH_3CN$ is 2 Da higher than that of the corresponding adducts produced without $NaBH_3CN$ in the incubation.

Experiments with bovine serum albumin (BSA) showed that adducts generated by both mechanisms are clearly visible even in the presence of an excess amount of unmodified BSA. For example, Figure 3a shows the original and the transformed mass spectrum generated by incubation of BSA with the acyl glucuronide of zomepirac. The signal at m/z 66,438 is due to BSA and the signal at m/z 66,652 is due to a common impurity in BSA (see reference 9 for more details). The signal at m/z 66,882 can be ascribed to the adduct formed by the imine mechanism. Remarkably, the measured mass of the adduct is only 6 Da (i.e. less than 0.01%) lower than the calculated mass. The signal at m/z 66,717 most likely corresponds with the nucleophilic displacement adduct. Addition of $NaBH_3CN$ to the incubation increases the abundance of the imine mechanism adduct significantly; see Figure 3b. In addition, peaks due to BSA molecules containing multiple adducts are visible as well. In contrast to BSA, other lysine-rich proteins such as cytochrome c, hemoglobin and myoglobin do not form any covalent adducts.

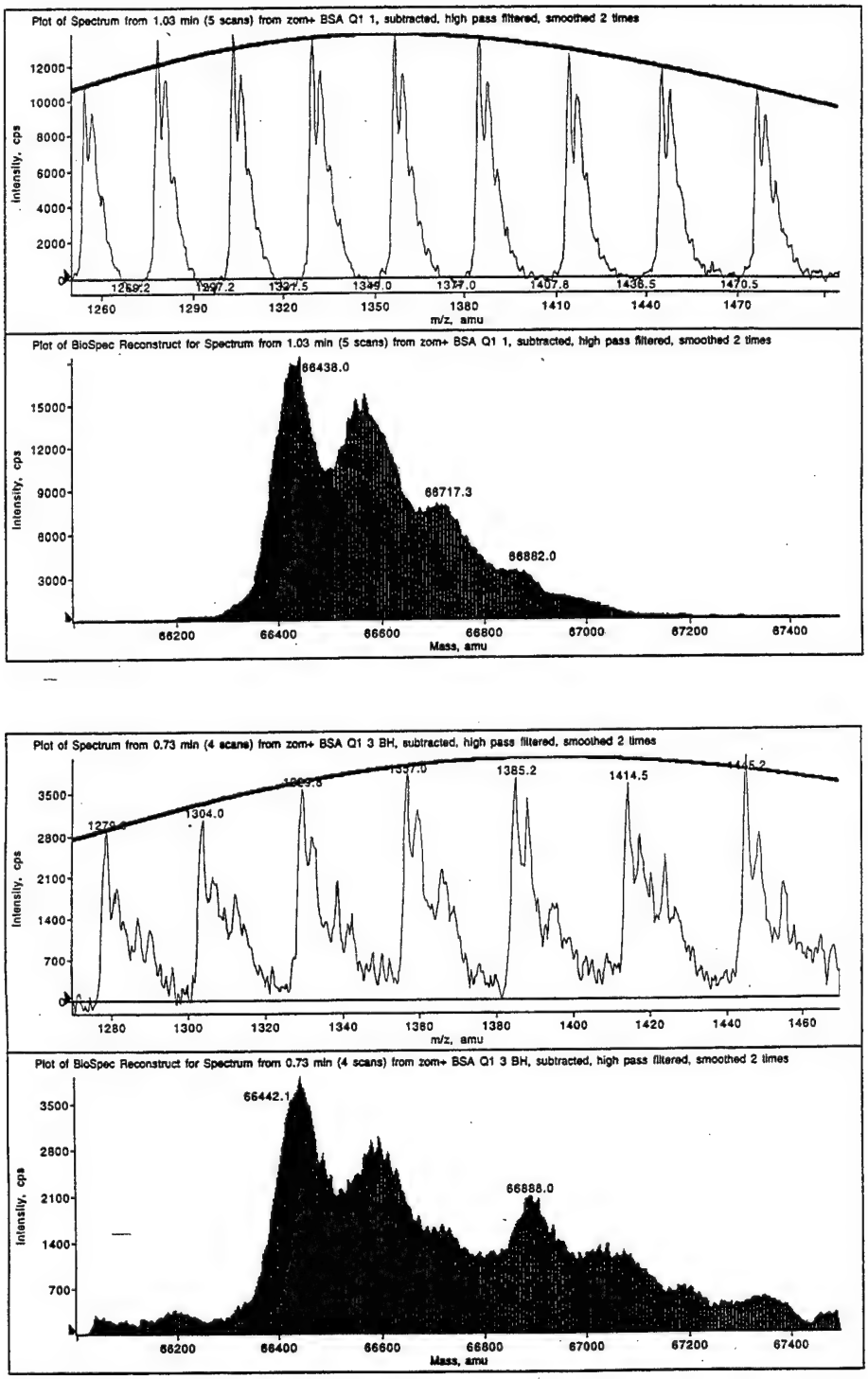


Figure 3. The original and the transformed mass spectrum for the bovine serum albumin + zomepirac acyl glucuronide incubation in the presence (b) and absence (a) of $NaBH_3CN$. See text for more details.

4. ACKNOWLEDGMENTS

The authors are grateful to Drs. A.Y.H. Lu and S.-H. L. Chiu for helpful suggestions.

5. REFERENCES

1. H.Spahn-Langguth and L.Z. Benet, *Drug Metab. Rev.*, **24**, 5-48 (1992).
2. M.J. Bailey and R.G. Dickinson, *Chem. Res. Toxicol.*, **9**, 659-666 (1996) and references therein.
3. P. Zia-Amirhosseini, J.C. Ojingwa, H. Spahn-Langguth, A.F. McDonagh and L.Z. Benet, *Clin. Pharmacol. Ther.*, **55**, 21-27 (1994).
4. L.T. Wade, J.G. Kenna and J. Caldwell, *Chem. Res. Toxicol.*, **10**, 546-555 (1997).
5. C. Fenselau, in "*Handbook of Experimental Pharmacology, Vol. 112*", ed. by F.C. Kauffman, Springer-Verlag, Berlin, Germany; pp 367-389.
6. S.J. Hargus, B.M. Martin, J.W. George and L.R. Pohl, *Chem. Res. Toxicol.*, **8**, 993-996 (1995).
7. (a) A. Ding, J.C. Ojingwa, A.F. McDonagh, A.L. Burlingame and L.Z. Benet, *Proc. Natl. Acad. Sci.*, **90**, 3797-3801 (1993). (b) A. Ding, P. Zia-Amirhosseini, A.F. McDonagh, A.L. Burlingame and L.Z. Benet, *Drug. Metab. Dispos.*, **23**, 369-376 (1995). (c) P. Zia-Amirhosseini, A. Ding, A.L. Burlingame, A.F. McDonagh and L.Z. Benet, *Biochem. J.*, **311**, 431-435 (1995).
8. H. Juteau, Y. Gareau and M. Labelle, *Tetrahedron Lett.*, **38**, 1481-1484 (1997).
9. L.Q. Huang, A. Paiva, R. Bhat and M. Wong, *J. Am. Soc. Mass Spectrom.*, **7**, 1219-1226 (1996).

Automated Collection of Fractions eluting from Capillary HPLC or CZE onto MALDI Targets and Membranes

Jean-Pierre Salzmann[1]), Remco van Soest[1]), Rudi Grimm[2]), Jean-Pierre Chervet [3])
[1]) LC Packings USA, San Francisco, California
[2]) Hewlett-Packard, Waldbronn, Germany
[3]) LC Packings, Amsterdam, The Netherlands

Introduction 1

MALDI-TOF/MS has gained wide acceptance for the analysis of proteins and peptides. These samples are often separated by miniaturized analytical techniques, i.e., Capillary LC and Capillary Electrophoresis (CZE).

At the small flow rates typical for these techniques, the volume of individual fractions is quite small — typically 0.1 to 2 µl — which makes manual collection and loading onto MALDI targets difficult.

In this poster we present a user programmable automated micro fraction collection system that is used with either Capillary LC or CZE: The effluent from the chromatographic system is automatically loaded onto the wells of the target. Due to the extremely high movement resolution of 2.5 micron, the sample is loaded very precisely within the area of the wells. Hence, scanning times can be reduced and loss of sample is prevented. A built-in dispensing system allows to add reagents or matrixes.

It is also possible to split and collect fractions in user definable proportions between two different substrates, e.g., MALDI-target and PVDF membrane.

The instrument is controlled by Windows software.

Experimental 2

Capillary LC separation

- µ-flows were achieved by using a conventional high pressure binary gradient system (HP 1100) and an Acurate™ microflow processor
- UV absorption detection was performed with a conventional HPLC detector (HP 1100 UV-VWD) equipped with an U-Z View™ Z-shaped flow cell

Capillary Electrophoresis separation

- a modular CE system controlled by a sample injection unit (Crystal) was used for the CE application
- UV absorption detection was performed with a conventional HPLC detector (ABI 785 A UV-VWD) equipped with an CE-Zell™ z-shaped flow cell

µ-Fraction collection

- µ-fraction collection was performed with a robot system (PROBOT) that collects fractions eluting from a µ-analytical systems

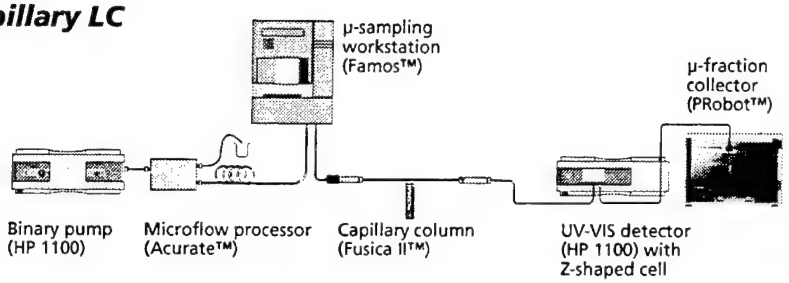
MALDI/TOF

- MALDI/TOF data was obtained with a series of different instruments (Bruker, PerSeptive Biosystems, Linear)

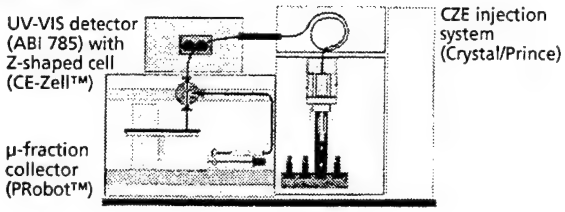
Experimental Set-up [1]

3

Capillary LC

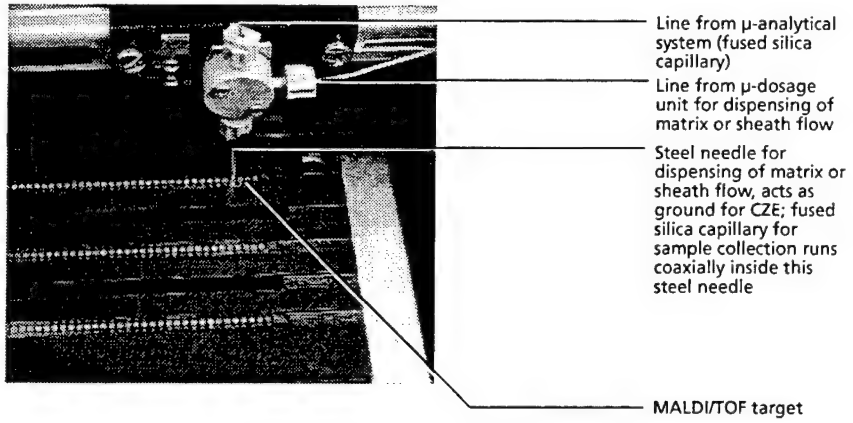


Capillary Electrophoresis



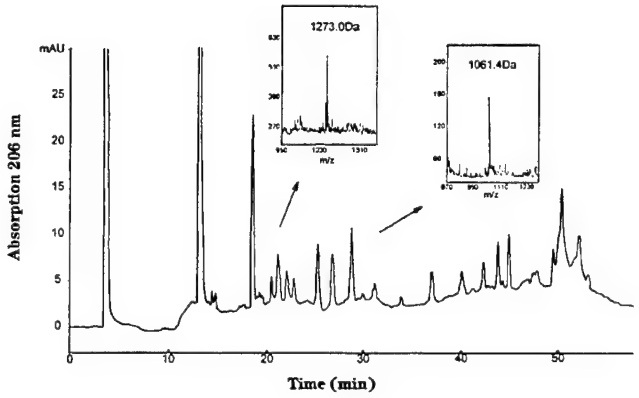
Experimental Set-up [2]

4



Results [1]

5

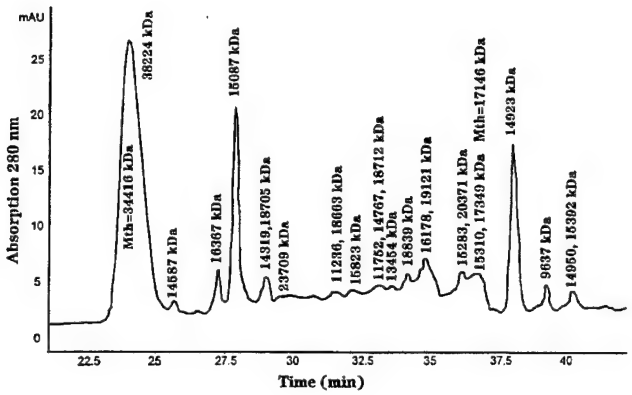


MALDI/TOF analysis of µ-fractions directly collected onto targets

Sample: tryptic digest of myoglobin (700 fmol)
Column: 300 µm x 25 cm C18 300 Å (Fusica II packed with Vydac)
Flow rate: 2 µl min
Fraction size: 1 µl (30 sec)

Results [2 a]

6

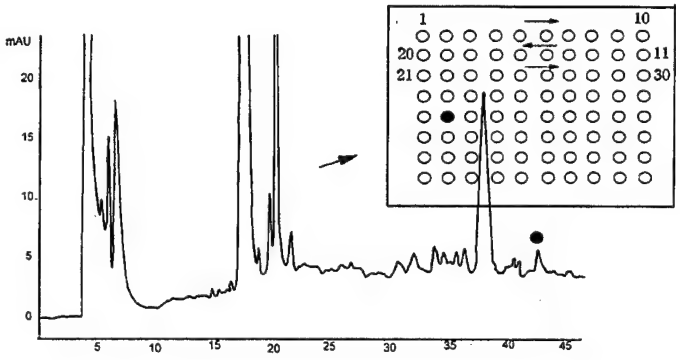


Automatic µ-fraction collection of a protein separation onto MALDI/TOF targets

Sample: 2% TCA protein extract of maize kernels
Column: 300 µm x 25 cm C18 300 Å (Fusica II packed with Vydac)
Flow rate: 3 µl min
Fraction size: 1.5 µl (30 sec)

Results [2 b]

7

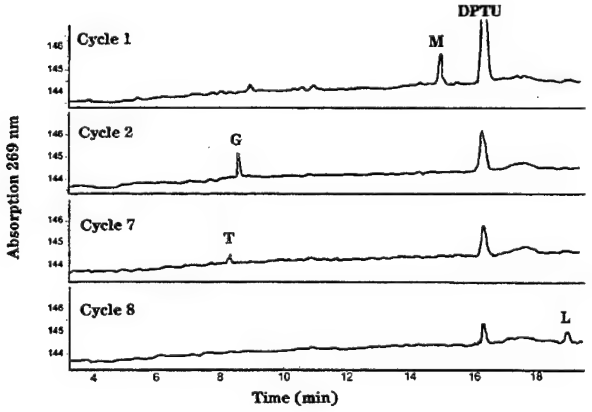


Automatic µ-fraction collection of a protein separation onto Selex-20 membrane for sequencing

Sample: 2% TCA protein extract of maize kernels
Column: 300 µm x 25 cm C18 300 Å (Fusica II packed with Vydac)
Flow rate: 3 µl min
Fraction size: 1.5 µl (30 sec)

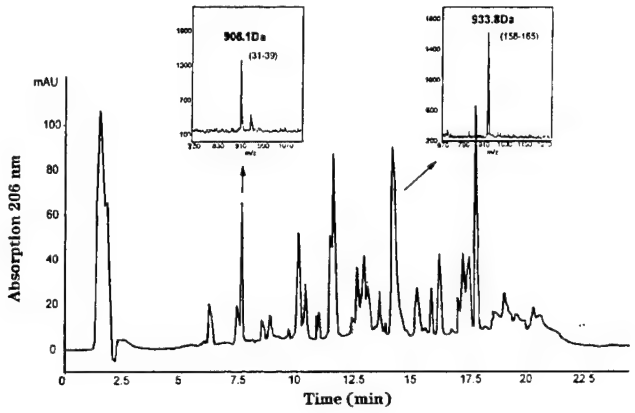
Results [2 c]

8



Protein sequence analysis of a peptide fraction on spot 42 (•) from maize protein automatically collected onto Selex-20 membrane
Sequencer: HP G1005A
Sequence: MQIFVKTLTGKTITLEVEE = Ubiquitin

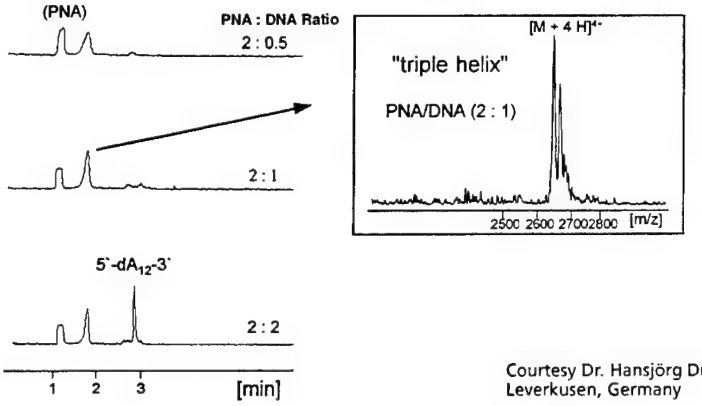
Results [3]



Peptide map followed by automatic µ-fraction collection onto MALDI/TOF targets

Sample: Tryptic digest of tobacco SSU
Column: 300 µm x 5 cm 1.5 µm C18 (Exmere)
Flow rate: 5 µl min
Fraction size: 2 µl (24 sec)

Results [4]



Courtesy Dr. Hansjörg Dürr, Bayer, Leverkusen, Germany

Hybridization assay with CE followed by MALDI/TOF with automated µ-fraction collection onto MALDI/TOF targets

Sample: PNA Peptide Nucleic Acid (T_{12}) + DNA
Buffer: Tris/borate (100 nmol *l $^{-1}$, pH 8.5, dextrane 60-90 (0.5.%)
Column: Fused silica capillary

Conclusions

- automated µ-fraction collection is fully compatible with Capillary LC and CZE
- automated collection onto MALDI/TOF targets
- matrix can be dispensed at time of collection or prior to collection
- due to small fraction volume drying is almost immediate
- fraction collection onto MALDI/TOF targets can be combined with collection onto other substrates, e.g., PVDF membranes

Capillary HPLC with Automated Fractionation into Nanospray Needles

Jean-Pierre Salzmann[1]), Remco van Soest[1]), Hans Vissers[2]), Ole Vorm[3]), Jean-Pierre Chervet[2])
[1]) LC Packings USA, San Francisco, California
[2]) LC Packings, Amsterdam, The Netherlands
[3]) The Protein Analysis Company, Odense, Denmark

Introduction 1

Glass capillaries with one end pulled to a very small orifice are used in Nano-Electrospray ionization sources to produce very small charged droplets. The liquid flow through the small orifice of the capillary is generated by the electrospray process and results in flow rates in the low nl/min range.

In this poster we describe a µ-sampling workstation that allows to automatically collect samples and fractions from Capillary LC into borosilicate Nano-Electrospray needles. Due to the high degree of automation, sample loss is greatly reduced, tedious and often laborious manual pipetting is circumvented, and the reproducibility of sample loading substantially increased.

The use of the µ-sampling workstation allows to perform complex automated sample clean-up steps prior to introducing the sample into the borosilicate needles. Examples will be shown that demonstrate the potential of automated sampling and µ-fractionation into Nano-Electrospray needles.

Experimental 2

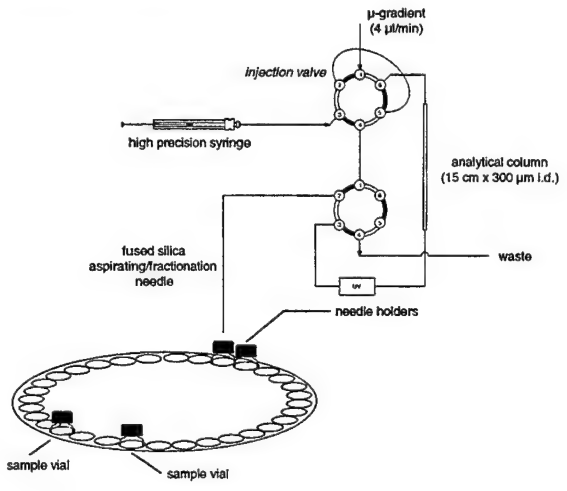
Capillary LC separation/µ-fractionation

- sample aspiration, injection and µ-fractionation were conducted fully automated on a FAMOS™ µ-sampling workstation
- µ-flows were achieved by using a conventional low pressure gradient system and an Acurate™ microflow processor
- UV absorption detection was performed with a conventional HPLC detector equipped with an U-Z View™ z-shaped flow cell

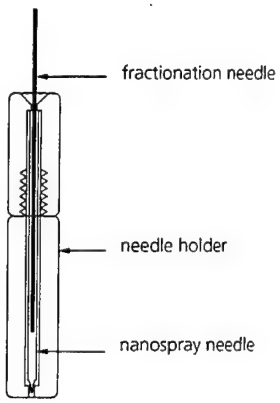
Nanospray MS

- nanospray ESI-MS spectra were recorded with a Finnigan LCQ mass spectrometer retrofitted with a nano ESI source head

Experimental set-up 3

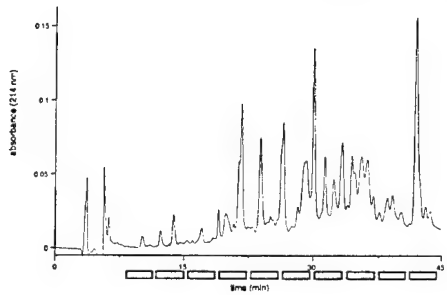


Needle Holder Design 4



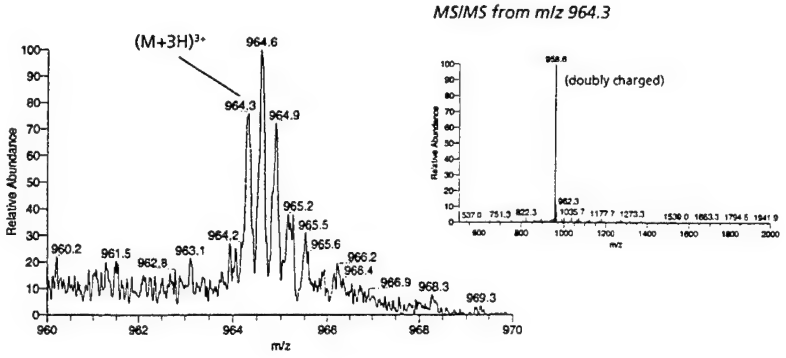
Chromatographic Results 5

sample:	tryptic digested fetuin
column:	15 cm x 300 µm I.D. C18 3 µm BDS
mobile phase:	A: 15 mM NH_4AC in CH_3CN/H_2O (5:95, v/v)
	B: 15 mM NH_4AC in CH_3CN/H_2O (80:20, v/v)
	5 - 50%B in 45 min
flow rate:	4 µl/min
detection:	UV absorption at 214 nm

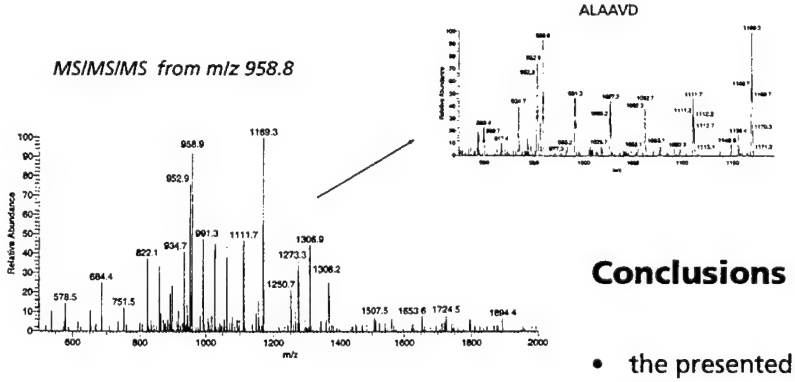


Nanospray ESI Specta 6

Nanospray spectra taken from one of the fractions



Nanospray MS/MS/MS sequencing 7



Conclusions 8

- the presented method allows for automated sampling and fractionation of minute samples into nanospray needles
- automated off-line technique: sample preparation, handling and transfer are eliminated
- sample enrichment and increased sample throughput feasible using automated microcolumn switching techniques
- nanospray suitable to sequence capillary LC purified peptides

INVESTIGATION OF THE MALDI-MS OF INTACT MICROBIAL CELLS AND THE FACTORS AFFECTING THEM

Karen L. Wahl, Nancy B. Valentine, Stephanie L. Gantt, Adam Saenz, Scott A. Clauss, and Mark T. Kingsley
Pacific Northwest National Laboratory
P. O. Box 999
Richland, WA 99352

ABSTRACT

Matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) is showing great potential for being a rapid screening tool for identifying microorganisms. Very reproducible MALDI mass spectra are obtained when the bacteria are grown under the same conditions, the bacterial sample is fresh, and the same MALDI matrix is used. Comparison of *Bacillus* strains by polymerase chain reaction (PCR) DNA fingerprinting and MALDI-MS obtained similar results indicating the ability of MALDI-MS to screen for minor differences in bacterial strains.

INTRODUCTION

The advent of matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) for biomolecule analysis has generated many powerful applications. The utility of this mass spectrometric technique for obtaining a unique fingerprint of ions from whole cell bacterial samples has been shown.[1,2] These exciting results show promise for using MALDI-MS as a rapid screening tool for bacterial identification from a wide variety of samples. The next step is to determine the ruggedness of the screening capability by investigating the parameters that influence the MALDI-MS data.

EXPERIMENTAL

A Kratos Kompact II MALDI time-of-flight mass spectrometer in the linear mode was used for all experiments. The mass axis was externally calibrated in positive ion mode with the singly and doubly protonated ions of cytochrome c and the matrix ions of sinapinic acid. Sinapinic acid, α-cyano-4-hydroxycinnamic acid (ACHC) and trifluoroacetic acid (TFA) were purchased from Aldrich (Gillingham-Dorset, United Kingdom and Milwaukee, WI, USA). The matrix solutions were 10 mg/mL in acetonitrile:water (1:1) with 0.1% TFA. Ammonium chloride and acetonitrile were obtained from J.T. Baker (Phillipsburg, NJ, USA). Water used was purified by a Milli-Q plus purification system (Millipore Corporation, Bedford, MA, USA).

Shewanella alga (BrY) was isolated from bottom sediments of the Great Bay Estuary, New Hampshire. The other bacterial samples were obtained from American Type Culture Collection (ATCC). In addition, a dry powder simulant was obtained from Dugway Proving Ground (DPG) and a "BG" spore preparation was obtained from Battelle Memorial Institute, Columbus Division (BMI). All samples were cultured in Difco tryptic soy broth (TSB). Before mass spectrometric analysis, the soy broth was washed from the cells with 2% ammonium chloride (NH_4Cl). A 1 mL aliquot of cells in TSB was centrifuged (12,000 rpm) to form a cell pellet. The supernatant was discarded, NH_4Cl was added and the pellet was resuspended by vortexing. The suspension was repelleted and washed twice. The final pellet was suspended in fresh 2% NH_4Cl solution and 1 µL was spotted onto the MALDI plate. The target spot was oven dried before adding 1 µL of sinapinic acid matrix and again oven dried. DNA from *Bacillus* cultures was subjected to PCR fingerprinting employing bacterial repeat DNA primer sets[3] (Table 1). Also DNA from 16 *Bacillus* cultures was subjected to PCR amplification by the Random Amplified Polymorphic DNA (RAPD) technique (Figure 3). The molecular weight standards included a 1kb DNA ladder and lambda *Hind*III digested DNA. The electrophoreses were run in a 1.4% agarose gel (0.5X TBE buffer)[4]. The RAPD primers were obtained from Operon Technologies.

RESULTS

The MALDI fingerprint spectra are reproducibly generated when the bacterial samples are in the same growth phase, and the matrix used for the MALDI analysis is the same. However, altering either culture age or matrix has a significant impact on the ions observed. For example, the gram-negative bacterium BrY has a very reproducible fingerprint of ions when a) the bacterial sample is fresh, b) the cells are in the log phase, and c) when sinapinic acid is used as the MALDI matrix. However, when the matrix is changed to α-cyano-4-hydroxycinnamic acid (ACHC), the ions observed vary from those obtained with sinapinic acid matrix (Figure 1). Further, more doubly-charged ions are observed for many of the bacteria when ACHC matrix is used. When the cell sample has aged, so that the cells are in death phase, the MALDI-MS spectra differ from those obtained with fresh cultures. Figure 2 shows the variation in ions observed when the washed bacterial sample has been stored in a freezer for two days. When the BrY cells completely lyse, the resulting MALDI mass spectrum is dominated by an ion at *m/z* 9420.

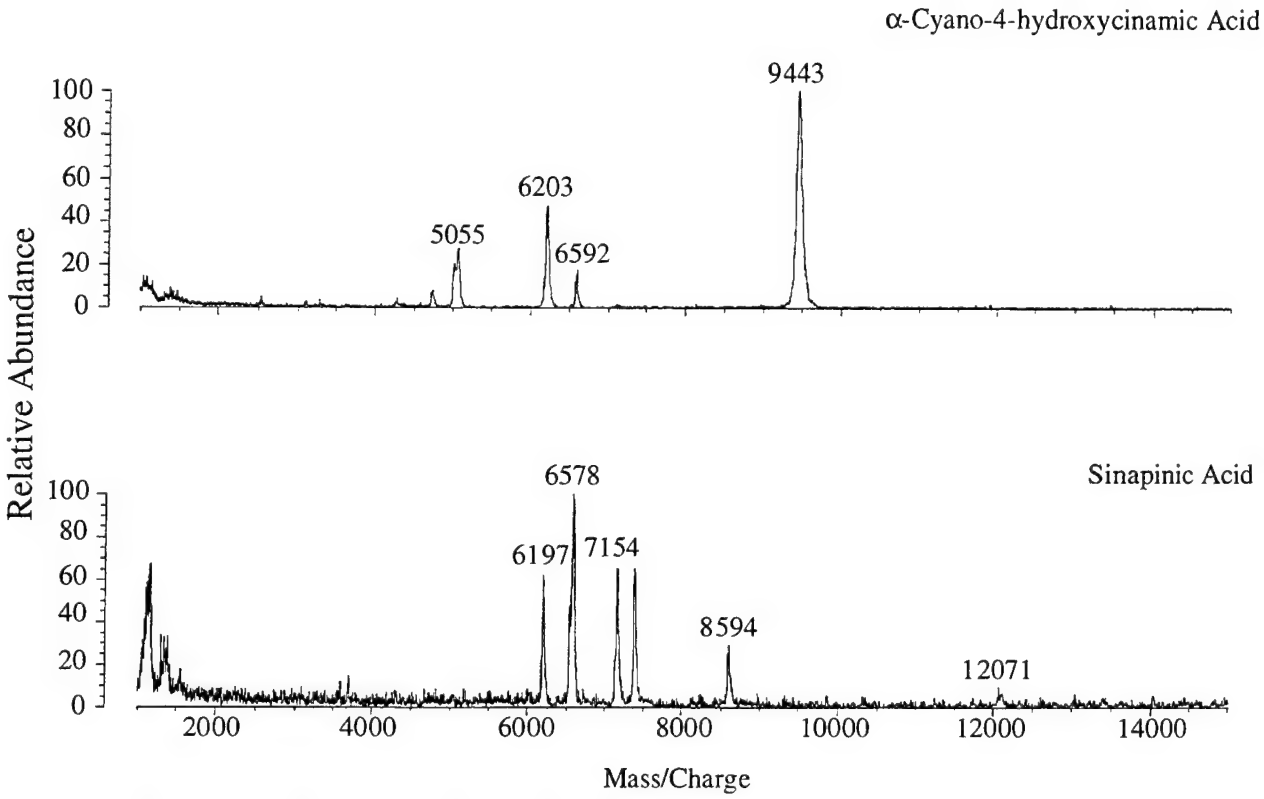


Figure 1. Influence of different MALDI matrices with BrY cells

A series of 13 *Bacillus* cultures were obtained from ATCC. Several of these cultures were assumed, based upon their ATCC catalog descriptions (orange colony color; formerly designated *Bacillus globigii* (BG)), to be similar to the "BG" simulant strain used, while others would serve as heterologous controls. An additional "BG" culture was obtained from BMI. This culture had been purified from the DPG dry simulant powder. These strains and their isolates are listed in Table 1. Upon initial laboratory culture, colony variants were observed (color, morphology) from most of the "BG" strains. Individual colony morphotypes were isolated and maintained as separate cultures. Cultures were grown and maintained on Difco trypticase soy medium. Genomic DNA was obtained by standard methods.

Analysis of these different cultures revealed that many were homologous, as evidenced by PCR gel traces shown in Figure 3. The results of these PCR gel traces are summarized in Table 1. Strains designated *B. atrophaeus* were similar to the "BG" isolates, but several bands in the DNA fingerprints were consistently different, regardless of the PCR primer sets employed. The MALDI-MS analyses of these bacterial samples are in agreement with the PCR results. The *Bacillus* strains that appear different based on PCR results, also produced different mass spectra by MALDI-MS.

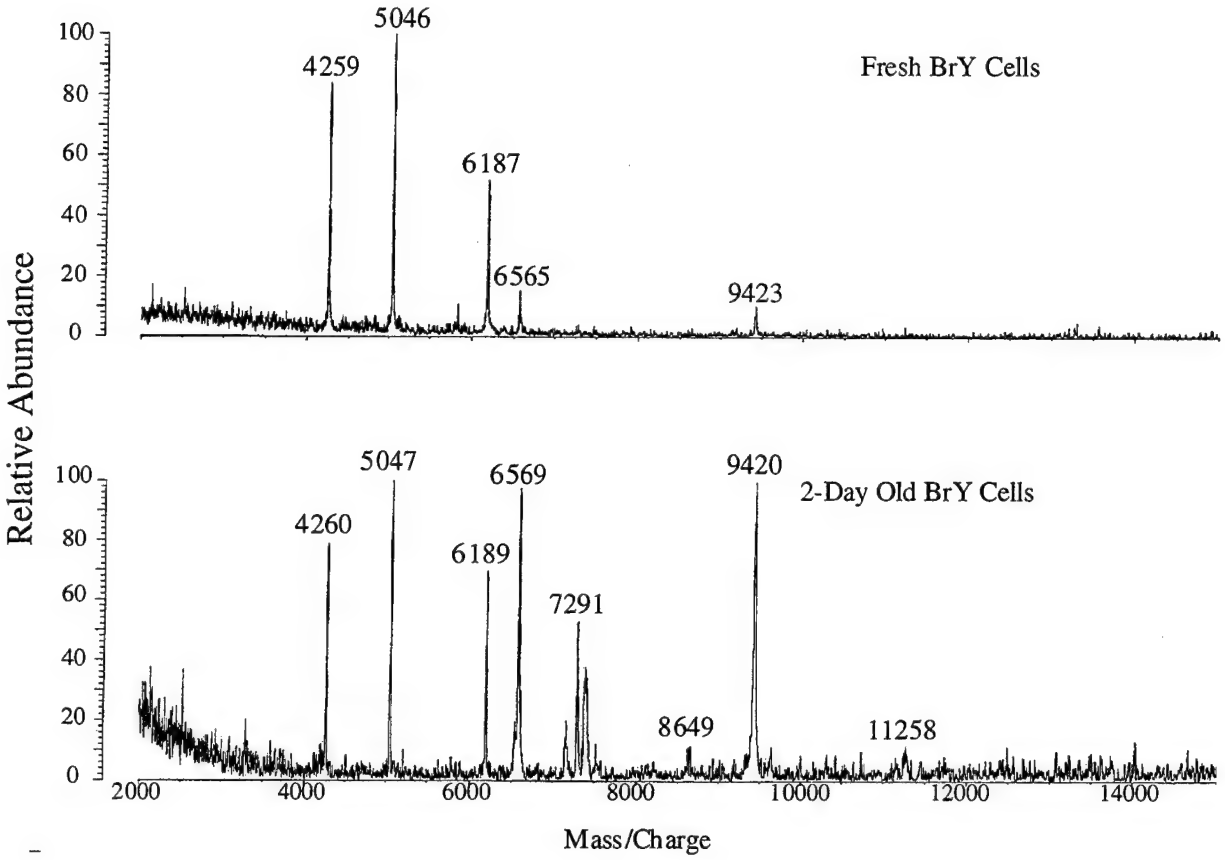


Figure 2. Influence of cell culture age on the ions observed by MALDI-MS

CONCLUSIONS

Matrix-assisted laser desorption/ionization mass spectrometry is a powerful rapid screening tool for microorganism identification. Reproducible mass spectral fingerprints are obtained from intact microbial cells with minimal sample handling. Several experimental variables, such as microbial cell growth conditions and analytical matrix selection, affect the mass spectral fingerprints. However, the MALDI mass spectra are reproducible under identical parameters. For example, variable cells with consistent analytical conditions yield reproducible MALDI-MS fingerprints. Aged cells with consistent analytical conditions also yield reproducible MALDI-MS fingerprints, but are different compared to viable cells. These variations may in fact provide valuable information to identify the origin and viability of the microbial sample. Cell integrity and age of culture are shown to affect the ions observed by MALDI-MS. The MALDI-MS results and the PCR nucleic acid determinations are in agreement for comparative analyses of bacterial cultures.

ACKNOWLEDGMENTS

Pacific Northwest National Laboratory is operated for the US Department of Energy by Battelle Memorial Institute under Contract DE-AC06-76RLO 1830. Funding was provided through Laboratory Directed Research and Development under the Detection and Characterization of Biological Pathogens Initiative. Technical assistance of Jeremy P. Green, Kelly Washington, and Justin Burns for PCR fingerprint data is gratefully acknowledged.

TABLE 1. PCR FINGERPRINTING OF "BG" STRAINS/REPEAT DNA PRIMER SETS

Strains		Primer Sets Rep[(1)(a)]	Eric[(a)]	Box[(a)]
ATCC 9372	(*B. subtilis*)	1	A	a
ATCC 31028	(*B. subtilis*)	1	A	a
ATCC 51189	(*B. subtilis*)	1	A	a
ATCC 49822	(*B. subtilis*, formerly *B. globigii*)	1	A	a
ATCC 49760	(*B. subtilis*, formerly *B. globigii*)	1	A	a
DPG BG	(orange from simulant powder)	1	A	a
BMI BG	(purified from DPG)	1	A	a
ATCC 7972	(*B. subtilis*)	1*	A*	a*
ATCC 6455	(*B. atrophaeus*, formerly *B.s.*)	1*	A*	a*
ATCC 49337	(*B. atrophaeus*, formerly *B.s.* var *niger*)	1*	A*	a*
ATCC 4516	(*B. circulans*)	2	B	b
ATCC 6501	(*B. subtilis*)	3	C	c
ATCC 14580	(*B. licheniformis*)	4	D	d
ATCC 14579	(*B. cereus*)	5	E	e
ATCC 14577	(*B. sphaericus*)	6	F	f
•ENVI	(white culture)	7	G	g

1 Patterns with same number or letter are identical
* indicates similar but slightly different (additional bands, etc.)
• different from "BG" and controls
(a)amplification protocols essentially as Louws et al., 1994[3]

ATTC - American Type Culture Collection
BMI - Battelle Memorial Institute
DPG - Dugway Proving Ground
ENVI - random environmental isolate

REFERENCES

1. Ross, PH and T Krishnamurthy. 1996. "Detection of Pathogenic and non-Pathogenic Bacteria by MALDI-MS Analysis of Protein Extracts and Whole Cell Suspensions." In Proceedings of the 44th ASMS Conference on Mass Spectrometry and Allied Topics, May 12-16, 1996, Portland, OR p 988.

2. Holland, RD, JG Wilkes, F Rafii, JB Sutherland, CC Persons, KJ Voorhees, and JO Lay, Jr. 1996. "Rapid Identification of Intact Whole Bacteria Based on Spectral Patterns using Matrix-assisted Laser Desorption/Ionization with Time-of-flight Mass Spectrometry." Rapid Commun. Mass Spectrom. **10**:1227-1232.

3. Louws, FJ, DW Fulbright, C Stephens, and FJ DeBruijin. 1994. "Specific Genomic Fingerprints of Phytophathogenic *Xanthomonas* and *Pseudomonas* Pathovars and Strains Generated with Repetitive Sequences and PCR." Appl. Environ. Microbiol. **60**(7):2286-2295.

4. Sambrook, J, EF Fritsch, and T Maniatis. 1989. "Molecular Cloning: A Laboratory Manual/Second Edition." Cold Spring Harbor Laboratory Press. Cold Spring Harbor, NY.

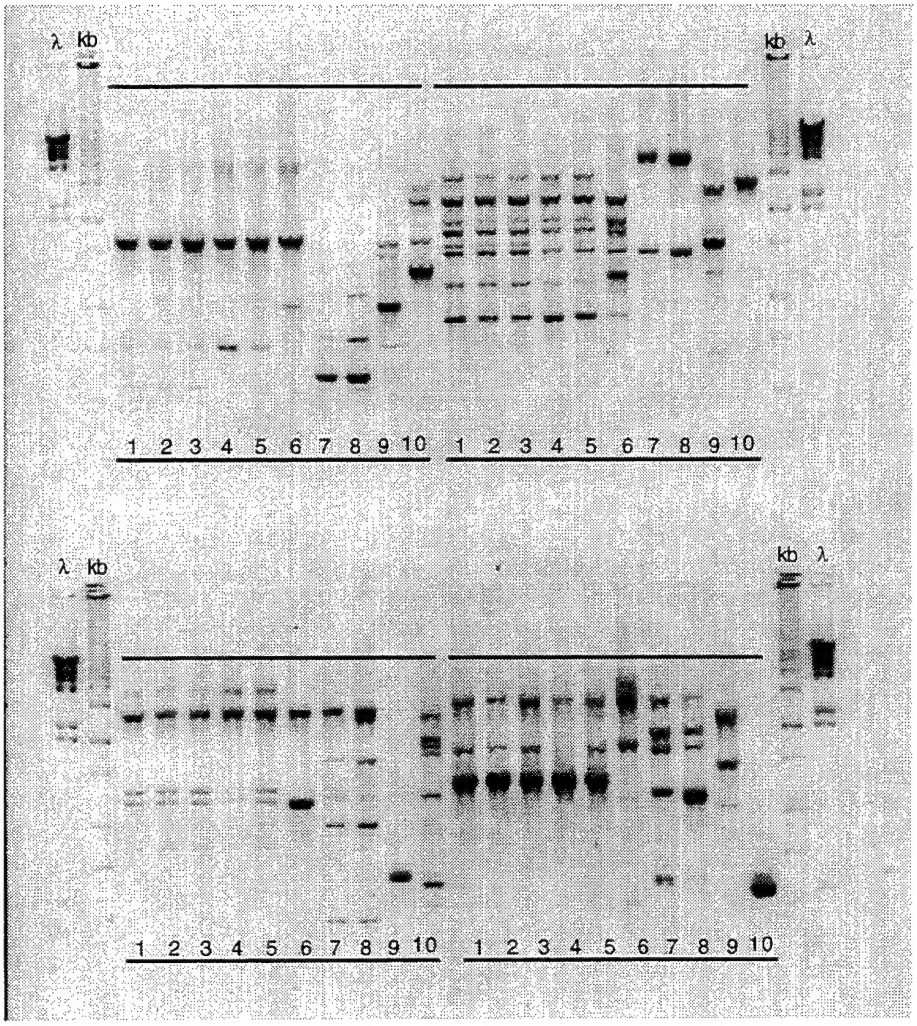


Figure 3. Analysis of *Bacillus* strains by four different RAPD primers. Lanes as follows: 1- ATCC 9372 (orange); 2 - ATCC 9372 (white); 3 - DPG BG (orange); 4 - BMI BG (orange); 5 - BMI BG (white); 6 - ATCC 49337; 7 - ENVI (white1); 8 - ENVI (white 2); 9 - ATCC 6501; 10 - ATCC 4516

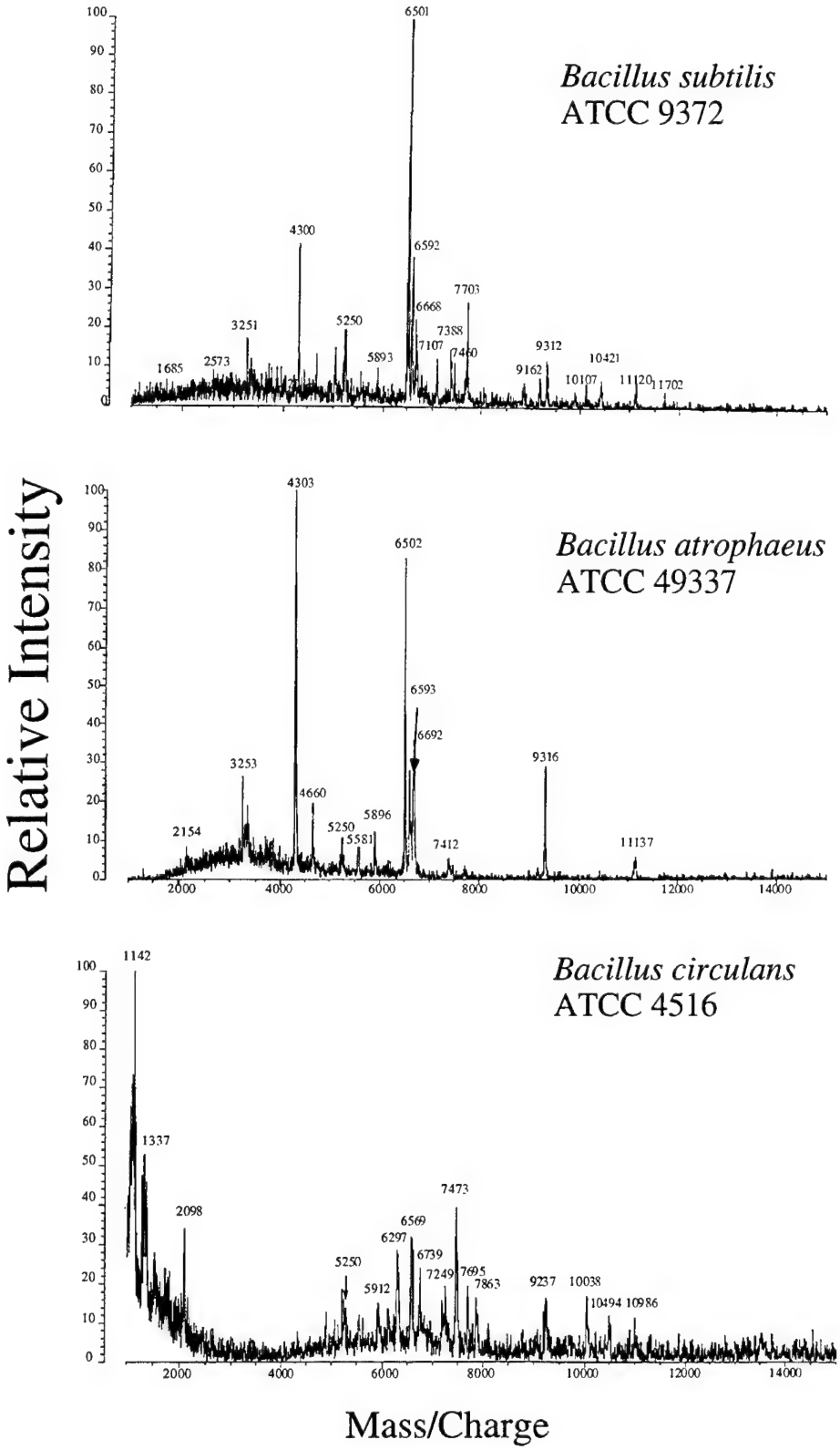


Figure 4. Comparison of three *Bacillus* samples by MALDI-MS

EFFECT OF ORGANIC BASES AND ACIDS ON BACTERIAL PHOSPHOLIPIDS DETERMINED BY ELECTROSPRAY MASS SPECTROMETRY

Michael W. Wensing, Karen S. Heroux
Geo-Centers, Inc.
Gunpowder Branch P.O. Box 68
Aberdeen Proving Ground, MD 21010

A. Peter Snyder
U.S. Army ERDEC, Attn: SCBRD-RTE
Aberdeen Proving Ground, MD 21010

ABSTRACT:

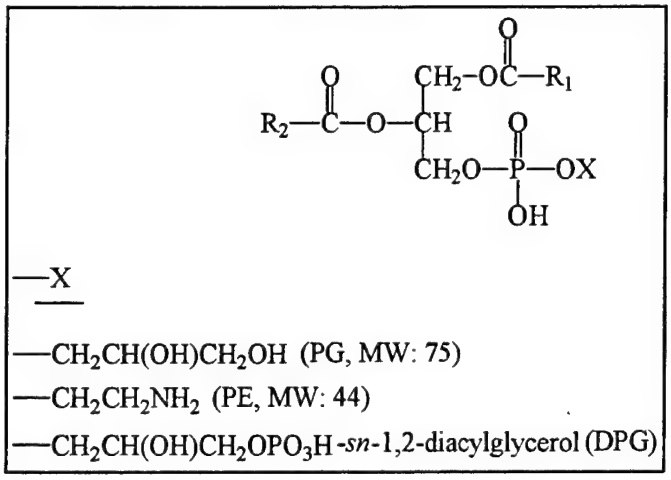
The effect of the addition of a base to bacterial extracts was studied to see if this would enhance the signal of diacylphosphatidylethanolamines (PEs). Phospholipids (PLs) have been used in bacterial profiling however, under negative ion ESI conditions the ionization efficiency of PLs is a strong function of the headgroup present on the PL. Those with glycerol headgroups are ionized most efficiently, i.e. diacylphosphatidylglycerols (PGs) and tetraacyldiphosphatidylglycerols (DPGs). However, PEs are not ionized efficiently when detected in nonpolar solvents, but are present in large quantities in several bacteria. A method is needed which can more sensitively detect PEs. The addition of piperidine to the mobile phase has shown that indeed, the signal of PEs can be enhanced greatly (i.e. 26 fold). This addition has increased the signal of PEs obtained from standard compounds as well as from bacterial extracts. This information was used to differentiate a limited set of bacteria.

INTRODUCTION:

Fast identification of bacterial species has proven to be an elusive goal. The method usually used, fatty acid methyl ester saponification (FAMES) requires extraction of the bacterial phospholipids, followed by base hydrolysis, and finally derivatization by methyl esterification so that the esterified fatty acids can be analyzed via gas chromatography. If the intact phospholipids could be analyzed directly, much time would be saved. Electrospray ionization mass spectrometry (ESI-MS) can analyze polar compounds such as phospholipids directly without derivatization unlike gas chromatography. However, the sensitivity of ESI is very dependent upon the headgroup (class) of the phospholipid. Diacylphosphatidylglycerols (PGs) and tetraacyldiphosphatidylglycerols have a high ionization efficiency, however, diacylphosphatidylethanolamines (PEs) (Scheme 1) have a low ionization efficiency. *Bacillus cereus*, a Gram positive bacterium, as well as Gram negative bacteria have a high percentage of PE, and the ability to detect PE as opposed to just PG may help differentiate these organisms from other closely related organisms.

EXPERIMENTAL:

Organic bases (piperidine and methylamine) were obtained from Aldrich (Milwaukee, WI). Phospholipid standards (99% purity, 97% positional purity) were obtained from Sigma (St. Louis, MO), and were used without further purification. Standard solutions were made by dissolving the phospholipids in 3:1 chloroform/methanol (v/v). Bacteria (American Type Culture Collection, Rockville, MD) were grown on trypticase soy agar (Becton Dickinson Microbiology Systems, Cockeysville, MD) plates for 24 hours. A loopful of bacteria was scooped from the plate, suspended in 500 μL 3:1 chloroform/methanol (v/v), vortexed for 1 minute to deagglomerate bacteria, 100 μL water was added, then the sample was centrifuged for 1 minute at 4000 rpm to aid in the separation of the phases. The chloroform layer was then removed for analysis by negative ion ESI-MS. Ten microliter aliquots of



Scheme 1 - Phospholipid structures by class.

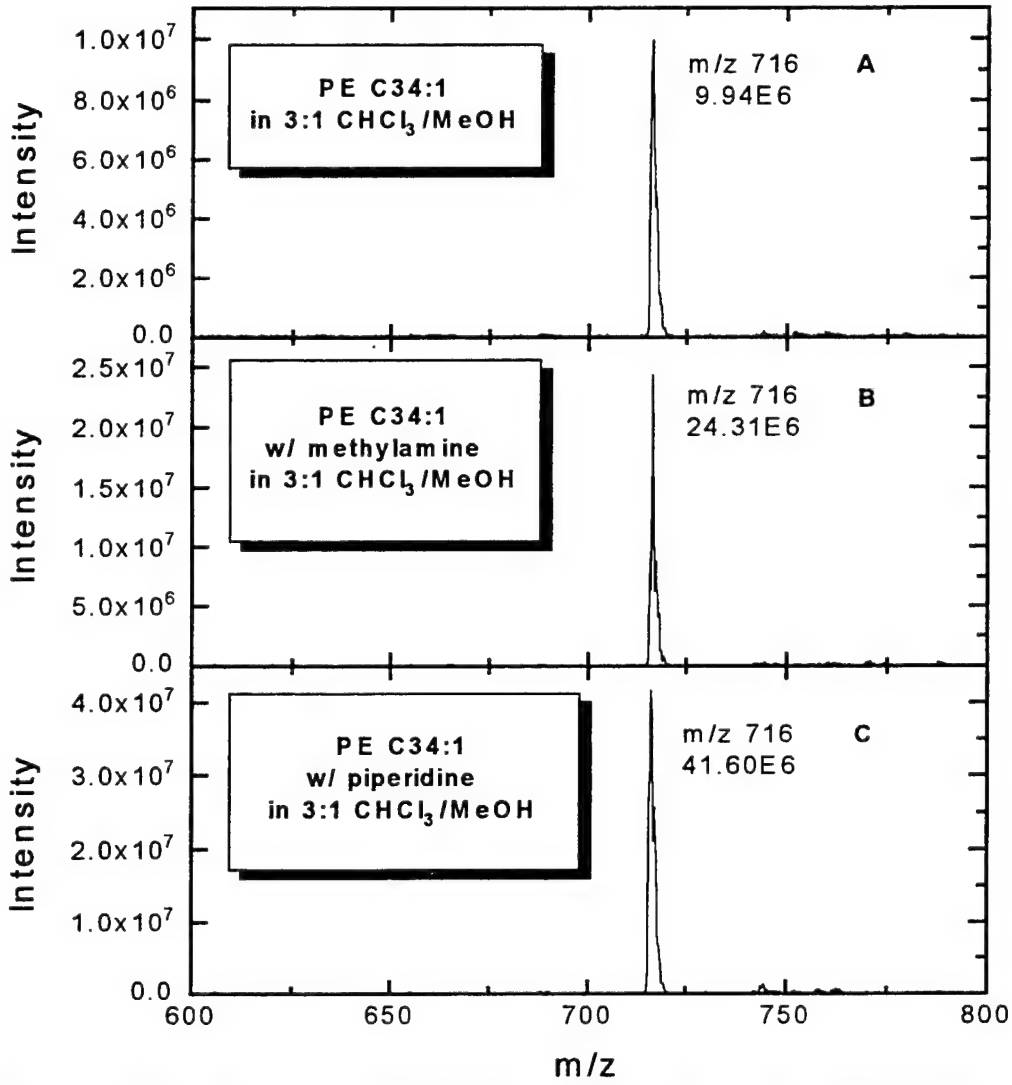


Figure 1 - Effect of a) no base b) methylamine and c) piperidine on the signal of a PE.

sample (either standard or extract) were swept into the ESI tandem mass spectrometer (Perkin-Elmer Sciex API III+, Concord, Ontario, Canada) at a flow rate of 50 μL/min using an HPLC pump (Waters 2690, Milford, MA).

RESULTS:

Standard compounds such as PE C34:1 (1-hexadecanoyl-2-[cis-9-octadecenoyl]-sn-glycero-3-phosphoethanolamine), where the number before the colon reflects the total number of carbons present in both fatty acid esters and the number after the colon reflects the total number of unsaturations present in total in both fatty acids, were dissolved in 3:1 chloroform/methanol and mixed with equal volumes of a) 3:1 chloroform/methanol, b) 0.5 mM methylamine (after dilution), and c) 0.5 mM piperidine (after dilution). Ten microliter aliquots were injected and detected by negative ion ESI-MS. The PE signal was enhanced with the addition of an organic base (Figure 1). This enhancement paralleled the base strength as determined by their K_b values. Methylamine is a weaker base than piperidine. Since piperidine provided a greater signal enhancement than methylamine for equivalent concentrations, piperidine was chosen for further studies. Enhancement of the PE signal (PE C28:0) was observed up to concentrations of 0.5 M piperidine, where a 26-fold gain in signal (Figure 2-top) was achieved with respect to a solution containing no base (Figure 2-bottom). One potential problem with modifiers is that analyte adducts with the modifer may be observed, such as formic acid adducts of PLs. An advantage of piperidine is that no piperidine adducts, i.e. M+85 peaks are observed even at high concentrations of piperidine.

Bacterial phospholipid extracts were prepared and analyzed with and without the addition of piperidine. Figure 3 shows the spectrum of a phospholipid extract of Gram negative *E. coli* 11775 which was obtained without piperidine. Only PGs are observed in great abundance when piperidine has not been added to the extract, specifically PG C32:1, PG C33:1, PG C34:1, PG C35:1, PG C36:2. PEs which should be present in the extract are not observed. With the addition of 0.5 M piperidine (Figure 3A) PEs can be observed in great abundance, specifically PE C32:1, PE C33:1, PE C34:1, PE C35:1 and PE C36:2.

Verification of the identity of these peaks was achieved by negative ion ESI-MS/MS (Figure 4A &4B). Phospholipids upon dissociation lose their fatty acids esters as carboxylate anions. The difference between the sum of these two fatty acids and the precursor ion is constant for a given class of phospholipid, 211 for PGs and 180 for PEs. The mass difference (211-180) corresponds to the difference in mass of the glycerol and ethanolamine headgroups. Figures 4A and 4B show the product ion mass spectra of PG C34:1 and PE C34:1 which were obtained from a phospholipid extract of *E. coli* 11775.

Figures 5-7 show the effect of piperidine on other phospholipid extracts obtained from Gram negative bacteria (*E. coli* 25922, *Brucella abortus*, *Yersinia rhodei* 43380, while Figure 8 shows the effect of piperidine on Gram positive *Bacillus cereus* 6464.

While addition of base allows one to obtain information on PEs as well as PGs, the resulting spectra are more complex. Quantitation is more complicated since PEs and PGs may be as little as 1 Da. apart, and due to the high number of carbons PEs and PGs have substantial (40-50%) M+1 isotope peaks. Therefore, isotope corrections must before applied before accurate quantitation can be achieved. In addition, the concentration of base must be consistent to determine the PE/PG ratio which can be useful in identifying a bacterium.

Detection limits of PEs were decreased 25 fold to 60 femtomoles with the addition of base, while little change was observed in the detection limits of PGs (30 femtomoles).

Table I lists the PG base peak and the PE base peak obtained from the ESI-MS mass spectra. These peaks may be used in a very simple scheme to classify the bacteria studied.

CONCLUSIONS:

The addition of an organic base is necessary to observe lipids of the phosphatidylethanolamine class. When 0.5 M piperidine was added a 25 fold decrease in the detection limit of PEs was observed, while no effect was observed in the detection limit of PGs. The addition of base resulted in complex ESI-MS spectra in the phospholipid extracts since PEs are observed 1 Da below PGs, and because the M+1 isotope peaks of PEs are substantial. Therefore, isotope correction is necessary before accurate quantitation can be achieved.

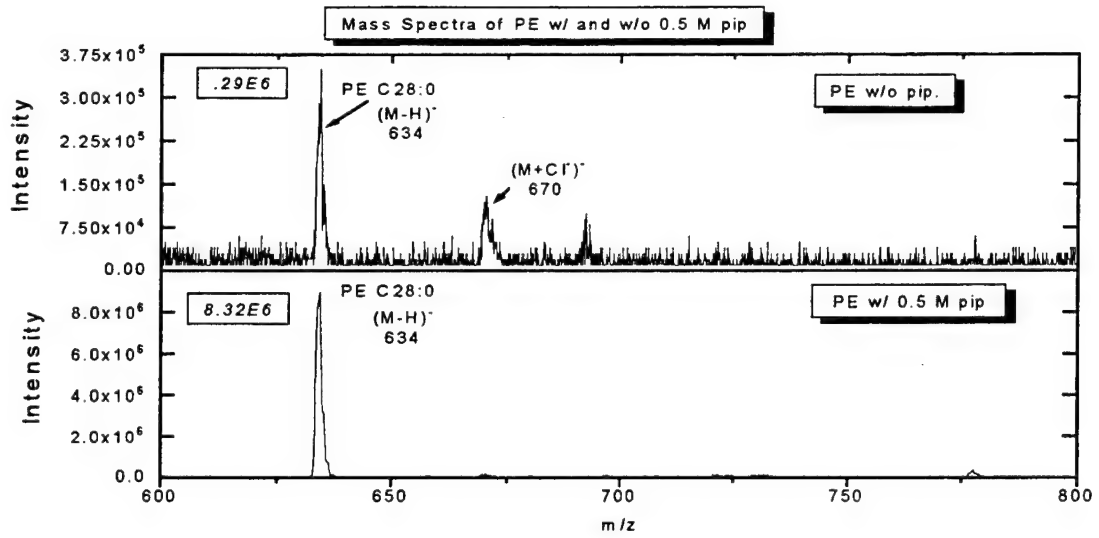


Figure 2- Dimyrisotyl PE C28:0 with (bottom) and without (top) piperidine.

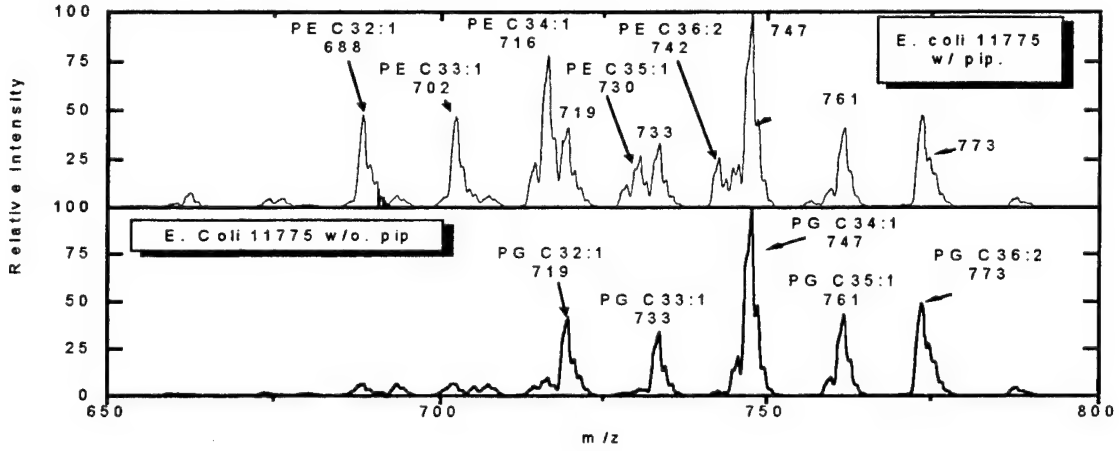


Figure 3 - *E. coli* 11775 with and without piperidine.

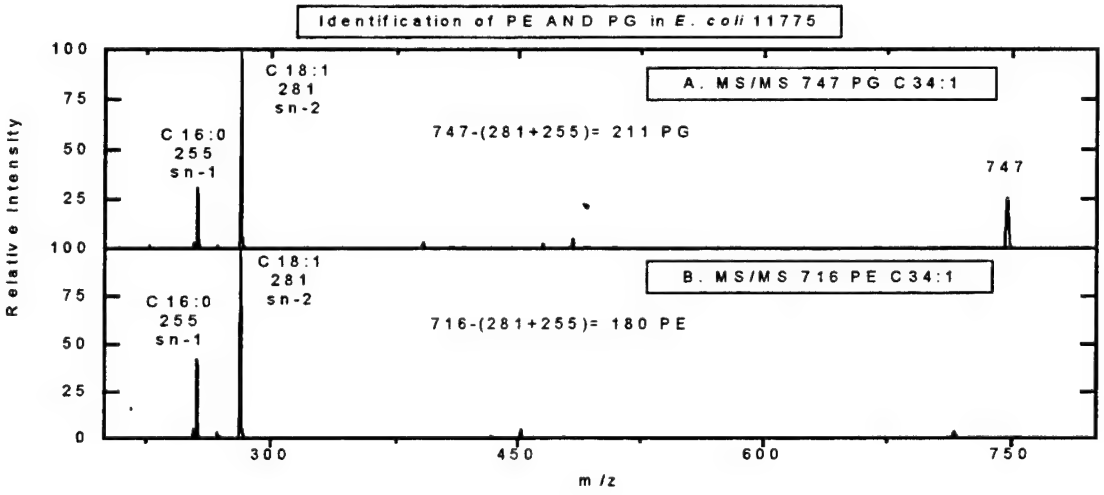


Figure 4 - Identification of a PE and PG in *E. coli* 11775.

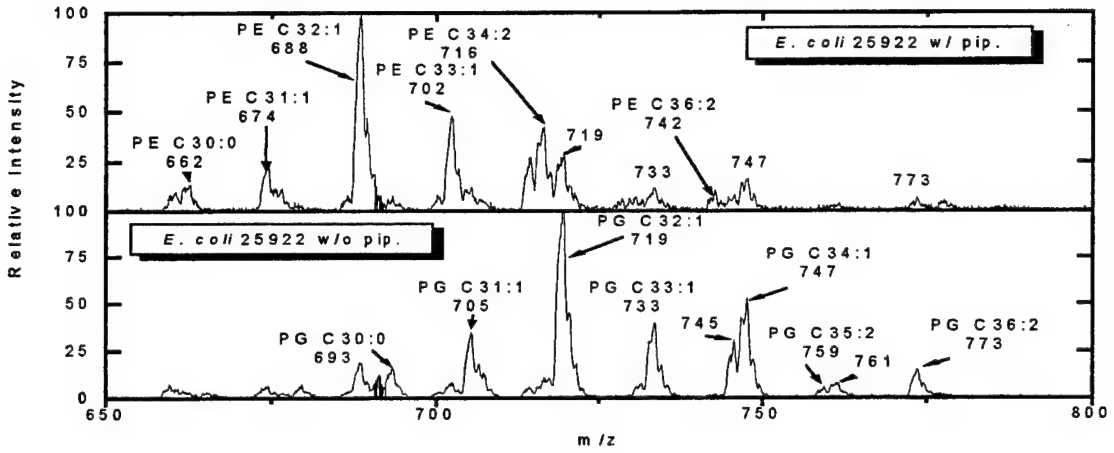


Figure 5 - *E. coli* 25922 with and without piperidine.

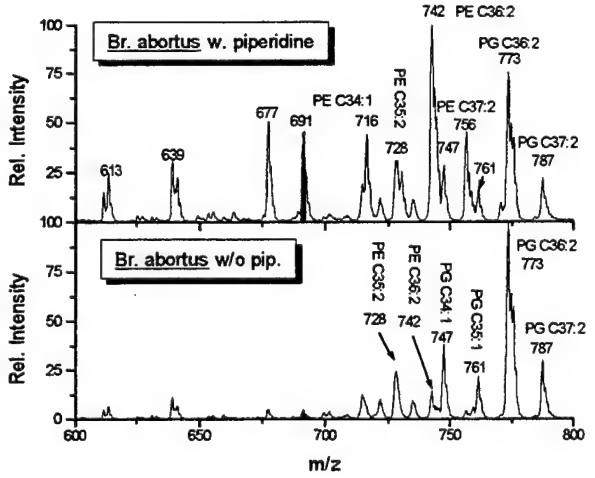


Figure 6 - *Brucella abortus* with and without piperidine.

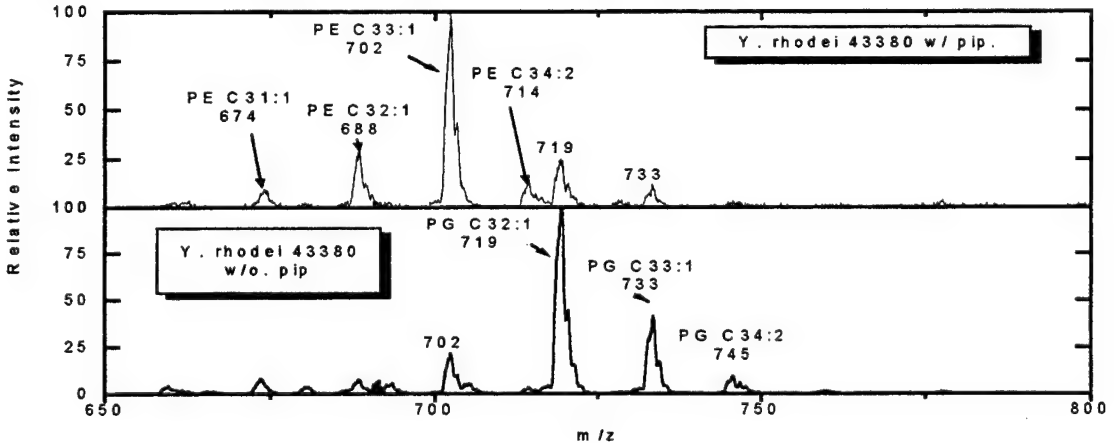


Figure 7 - *Yersinia rhodei* 43380 with and without piperidine.

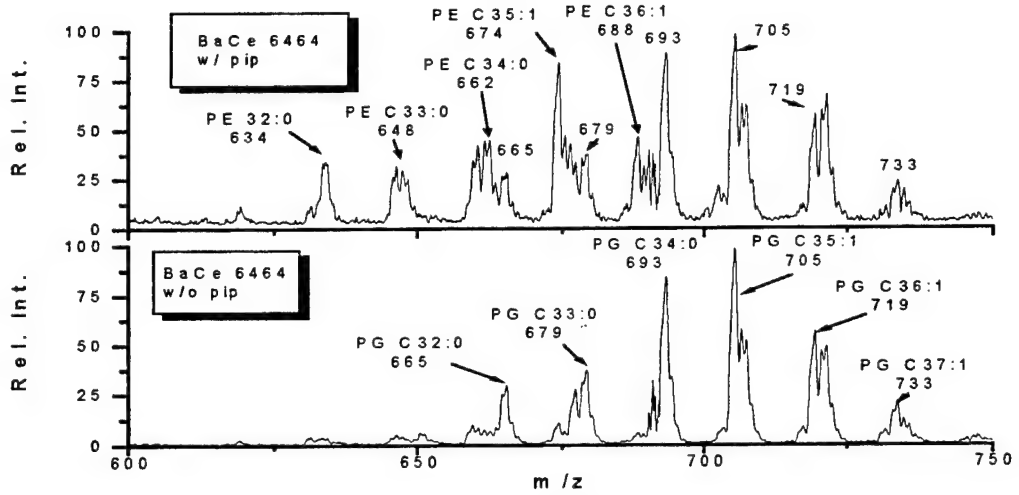


Figure 8 - *Bacillus cereus* 6464 with and without piperidine.

TABLE I - Most Intense PE and PG peaks present in the Bacterial Extracts

	PE C31:1 674	PE C32:1 688	PE C33:1 702	PG C31:1 705	PE C34:1 716	PG C32:1 719	PE C36:2 742	PG C34:1 747	PG C36:2 773
Y. rhodei 43380			X			X			
E. coli 11775					X			X	
E. coli 25922		X				X			
B. cereus 6464	X			X					
Br. abortus							X		X

REAL-TIME MASS SPECTROMETRY OF INDIVIDUAL AIRBORNE BACTERIA

William B. Whitten, Rainer A. Gieray, Peter T. A. Reilly, Mo Yang, and J. Michael Ramsey
Oak Ridge National Laboratory
Oak Ridge, Tennessee 37831-6142 USA

ABSTRACT

A method for real-time detection of individual bacteria is described. Airborne bacteria and bacterial spores are directly sampled by laser ablation in an ion trap mass spectrometer with an atmospheric pressure inlet system. Either positive or negative ion mass spectra can be obtained. Ions of a particular value of m/z can be further characterized by tandem mass spectrometry in the ion trap. Spectra averaged from several hundred individual bacteria of the same species appear to differ somewhat from spectra of bacteria of other species and to be readily distinguishable from spectra of nonbiological particles.

1. INTRODUCTION

We are developing a method for real-time analysis of airborne microparticles based on laser ablation in an ion trap mass spectrometer.[1-3] Experiments with aerosolized bacteria show some promise for discrimination of bacteria from particles of nonbiological origin. While there are large fluctuations from cell to cell, making identification of individual organisms tenuous, spectra averaged over several hundred cells differ slightly from species to species. The present approach might be used to provide an advanced warning of substantial changes in the background level of airborne bacteria or bacterial spores, triggering a more specific but labor and time-intensive identification process.

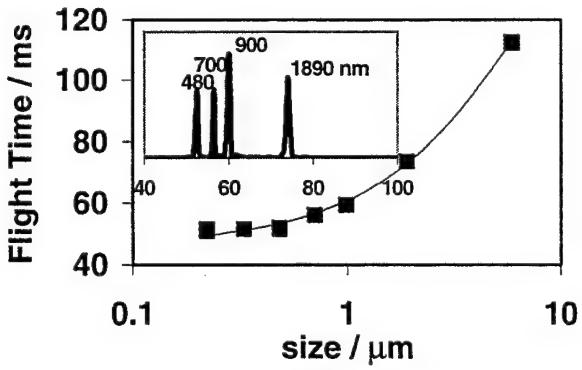


Fig. 2

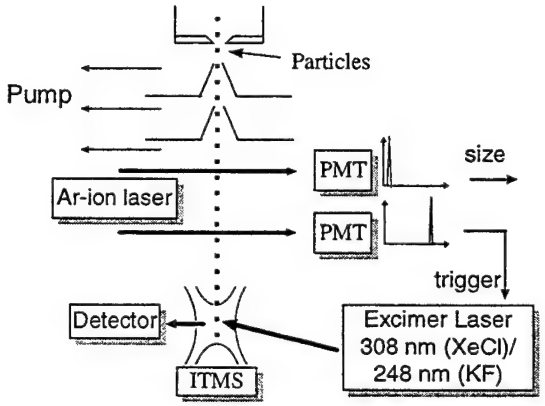


Fig. 1

The experimental approach is diagrammed in Fig. 1. Airborne particles enter the apparatus through an atmospheric orifice and are isolated from the surrounding air by skimmers. The particles are individually detected as they pass through two CW laser beams. The two timing pulses are used to trigger a pulsed excimer laser that samples and generates ions from the particle within the electrodes of an ion trap mass spectrometer. After the laser pulse, the stored ions are mass analyzed by conventional ion trap methods.[4]

The time for a particle to pass through the two CW laser beams is a function of the aerodynamic size of the particle, as shown in Fig. 2. A calibration curve is obtained from microspheres of known size. For particles of micrometer dimensions, we can obtain a reliable estimate of particle size in this way. The digitized transit time for each particle can be stored together with its mass spectrum. Knowledge of the particle size provides important additional information for particle classification.

A large number of ions can be produced from a single microorganism - enough to fill the ion trap. An example of a single particle mass spectrum is shown in Fig. 3. This is a positive-ion mass spectrum of a single *Bacillus subtilis* cell, obtained with a laser pulse of approximately 5 mJ at 308 nm wavelength, focused into a spot of 0.5 mm diameter. The mass scan was initiated at 50 Da to avoid detector overload from the lighter ions, primarily potassium ions at m/z 39. A negative ion mass spectrum of a single cell of the same species is shown in Fig. 4. We have also shown in this figure some of the ions that have been studied by tandem mass spectrometry with arrows indicating the fragment ions that were detected in collisionally-induced dissociation experiments. Many of the prominent ions in these spectra appear to contain phosphate or potassium. With UV laser ablation, we have not been successful in generating appreciable numbers of ions with m/z greater than 300 Da. We assume that the ions we detect are mostly fragments of the large molecular constituents of the cell membranes disrupted by the high intensity laser pulses.[5]

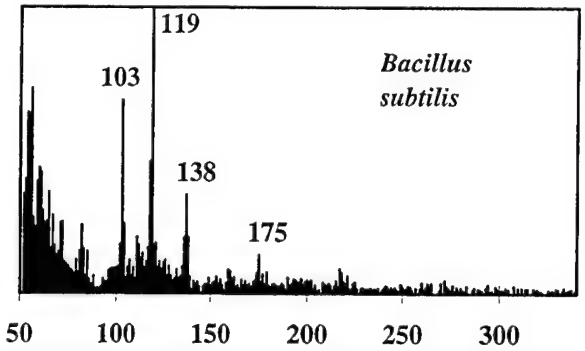


Fig. 3

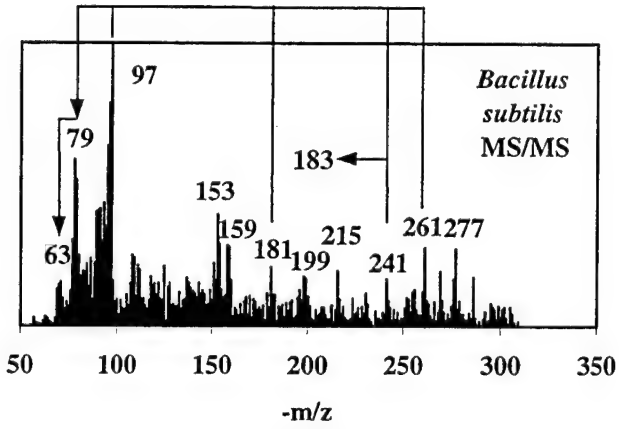


Fig. 4

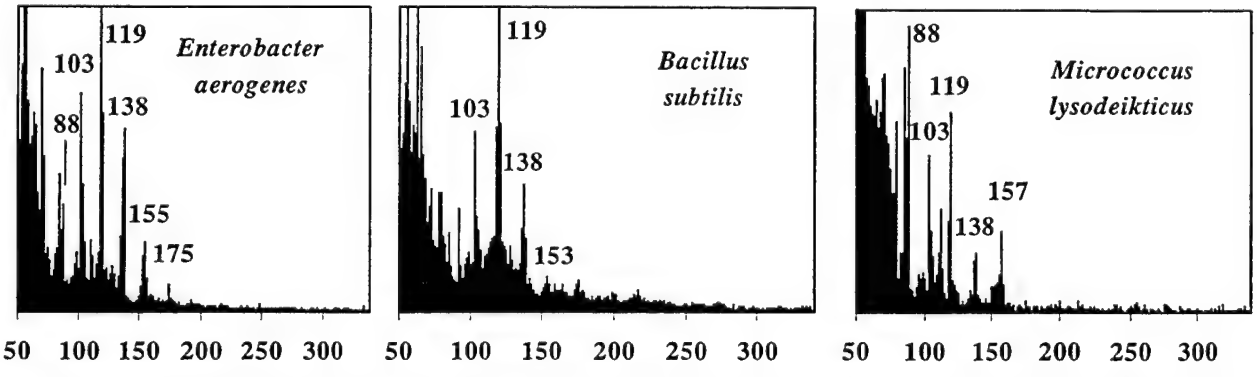


Fig. 5. a., Averaged single cell positive ion mass spectra of three species, *E. aerogenes, B. subtilis,* and *M. lysodeikticus.*

Positive ion mass spectra are shown in Fig. 5 for three species of bacteria together with spectra of three types of nonbiological particles that might be encountered in the environment. Because of the large

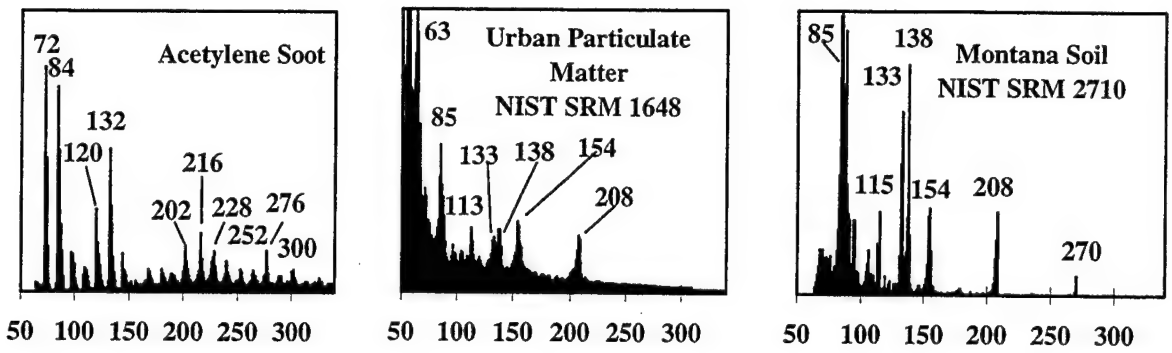


Fig. 5 b., averaged mass spectra of nonbiological particles, acetylene soot, urban airborne particulate matter (NIST SRM 1648), and Montana soil (NIST SRM 2710).

variation in the spectra of individual cells, even of the same species, we have presented spectra averaged over several hundred single particle measurements. The similarity of the bacterial mass spectra and the differences from the nonbacterial particles are apparent. We have teamed with John S. Wagner's group at Sandia National Laboratory to explore statistical discrimination of individual particle mass spectra.

In summary, a method for the real-time detection of airborne bacteria by laser ablation mass spectrometry in an ion trap has been developed. Microparticles are sampled directly from the air by a particle inlet system into the vacuum chamber of a mass spectrometer. An incoming particle is detected as it passes through two CW laser beams and a pulsed laser is triggered to intercept the particle for laser ablation/ionization and subsequent mass analysis in the ion trap mass spectrometer. Either positive or negative ions can be studied and ions of a particular value of m/z can be further characterized by tandem mass spectrometry in the ion trap. Statistical methods to discriminate between bacteria and other airborne particles appear feasible. We are currently exploring alternative means of sampling and ionization to extend the mass range to higher values of m/z.

2. ACKNOWLEDGEMENT

This research was sponsored by the Office of Research and Development, U.S. Department of Energy, under contract DE-AC05-96OR22464 with Oak Ridge National Laboratory managed by Lockheeed Martin Energy Research Corp.

3. REFERENCES

1. M. P. Sinha, in Instrumental Methods for Rapid Microbiological Analysis, W. H. Nelson, Ed., VCH Publishers, Inc., U.S.A., 1985, pp 165-192.
2. P. T. A. Reilly, R. A. Gieray, M. Yang, W. B. Whitten, and J. M. Ramsey, Anal. Chem. **69**, 36-39 (1997).
3. R. A. Gieray, P. T. A. Reilly, M. Yang, W. B. Whitten, and J. M. Ramsey, J. Microbiol. Methods (in press).
4. Practical Aspects of Ion Trap Mass Spectrometry, vol I, R. E. March and J. F. J. Todd, Eds., CRC Press, Boca Raton, 1995.
5. D. N. Heller, C. Fenselau, R. J. Cotter, P. Demirev, J. K. Olthoff, J. Honovich, M. Uy, T. Tanaka, and Y. Kishimoto, Biochem. Biophys. Res. Commun. **142**, 194-199 (1987).

Blank

UNAVAILABLE ABSTRACTS / MANUSCRIPTS

Genes, Genomes and Systems: New Opportunities for Mass Spectrometry
Prof. Leroy Hood, University of Washington, Seattle, WA

Mass Spectrometry in Viral Analysis
Dr. Gary Siuzdak, Scripps Research Institute, La Jolla, CA

Time-of-Flight Mass Spectrometry: New Technology for Biological Research and Agent Detection
Prof. Robert Cotter, Johns Hopkins University School of Medicine, Baltimore, MD

New Developments in MS/MS Ion-Trap Technology
Dr. Ian Jardine, Finnigan-MAT Corporation, San Jose, CA

High Performance Time-of-Flight Mass Spectrometry for Biological Analysis
Dr. Brian Musselman, PerSeptive Biosystems, Inc., Framingham, MA

Electrospray Mass Spectrometry for Detection and Characterisation of Intact Viruses
Dominique Despeyroux, PLSD, CBD Porton Down, Salisbury, United Kingdom

A Novel Dual-Dialysis Approach for Fast and Efficient Sample On-Line Fractionation and Clean-up for ESI-MS Analysis of Biological Agents
Chuanliang Liu, Pacific Northwest National Laboratory, Richland, WA

A Novel RF-Ion Funnel/Trap for Improved Sensitivity in the Detection of Biological Agents by Mass Spectrometry
Scott Shaffer, Pacific Northwest National Laboratory, Richland, WA

Chemical Biological Mass Spectrometer (CBMS)
David W. Sickenberger, U.S. Army Edgewood Research, Development and Engineering Center, Aberdeen Proving Ground, MD

Bacteria Detection and Differentiation with *In Situ* Thermal Hydrolysis-Methylation and the Chemical/Biological Mass Spectrometer (CBMS)
Kent J. Voorhees, Colorado School of Mines, Golden, CO

AUTHOR INDEX

AUTHOR INDEX

AUTHOR INDEX

Blank

PARTICIPANTS

- A -

Dr. Ruedi **Aebersold**
Dr. Sunia **Afzaal**
Dr. Philip C. **Andrews**

- B -

Dr. Ruth **Barak**
Douglas F. **Barofsky**
Stacy-Ann **Barshick**
Dr. Dipak K. **Basu**
Dr. Christopher **Becker**
Dr. Peter **Bel**
Klaus **Biemann**
Dr. Gavin E. **Black**
William C. **Blanchard**
Dr. Adam **Brockman**
Dr. Robert **Burkhalter**

- C -

Dr. John N. **Callahan**
Dwight A. **Clayton**
R. Graham **Cooks**
Charles **Coonley**
Dr. Timothy J. **Cornish**
Catherine E. **Costello**
Robert J. **Cotter**

- D -

Dr. Dominique **Despeyroux-Hill**
BG John C. **Doesburg**
Roma **Dogra**
Dr. Mildred A. **Donlon**
Dr. Vladimir **Doroshenko**
Thomas **Driscoll**
Dr. Ytzhak **Duer**

- E -

Dr. Michael **Easterling**
Dr. Charles G. **Edmonds**
Dr. Maria L. **Embrey**
Dr. John W. **Ezzell**

- F -

Dr. Catherine C. **Fenselau**
Alvin **Fox**

- G -

Stuart A. **Gaines**
Bradford W. **Gibson**
Dr. Kevin **Giles**
Dr. Wayne H. **Griest**
Dr. Harold **Guard**

- H -

Dr. Kevin J. **Hart**
David **Heller**
Karen **Heroux**
Franz **Hillenkamp**
Dr. Steven A. **Hofstadler**
Dr. J. Jerome **Holton**
Leroy **Hood**
Dr. Cornelis **Hop**
Donald F. **Hunt**

- J -

Dr. E.M. **Jakubowski**
Dr. Ian **Jardine**

PARTICIPANTS

- K -

Dr. Surinder **Kaur**
Anne **Keleher**
Dr. Avi **Keysary**
Dr. Slava **Kovtoun**
Dr. Thaiya **Krishnamurthy**

- L -

Dr. Stephen A. **Lammert**
Dr. Terry D. **Lee**
Liang **Li**
Dr. Chuanliang **Liu**
Dr. S. Randolph **Long**
Dr. Joseph A. **Loo**
David M. **Lubman**

- M -

Dr. Raymond **Mariella**
Dr. Scott A. **McLuckey**
Sebastian **Meyer-Plath**
Dr. Mike **Morris**
Dr. Brian D. **Musselman**

- O -

Dr. Rachel R. **Ogorzalek Loo**
Dr. Susan **Ohorodnik**

- P -

D.G. **Parekh**
Maxine **Pennington**
Kirkman R. **Phelps**
Dr. Elvis O. **Price**
Barry F. **Pritchard**

- Q -

Dr. Jun **Qin**

- R -

Uma **Rajamani**
Dr. Uwe **Rapp**
Vernon N. **Reinhold**
Dr. Drew C. **Richardson**
Carol **Robinson**
Dr. Dennis **Roser**
Dr. Philip L. **Ross**

- S -

Dr. Jean-Pierre **Salzmann**
Dr. Jeffrey **Shabanowitz**
David W. **Sickenberger**
Dr. Gary **Siuzdak**
J. Richard **Smith**
Dr. Richard D. **Smith**
Dr. A. Peter **Snyder**
Dr. Jhobe **Steadman**
Dr. James L. **Stephenson**, Jr.

- V -

Dr. Arpad **Vass**
Dr. Peter **Verhaert**
Dr. Kent J. **Voorhees**

- W -

Dr. Karen L. **Wahl**
Shengjun **Wang**
Dr. Michael W. **Wensing**
Dr. Donald L. **Wenzlick**
David C. **White**
Dr. William B. **Whitten**
Russell **Williams**
Dr. Dennis A. **Wolf**

PARTICIPANTS

Dr. Ngai **Wong**
Dr. Amina S. **Woods**
James R. **Wright**
Dr. John **Wronka**

- Y -

John R. **Yates**

- Z -

CPT Andrew T. **Zander**